

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:34:09 ; Search time 7963.34 Seconds  
(without alignments)  
4867.828 Million cell updates/sec

Title: US-10-051-909-32  
Perfect score: 4075  
Sequence: 1 IRSGSLAVQTFTPLDRR.....PLEVITEFPAVGAQQAAXA 800

Scoring table:  
BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool.p/US10051909/runat.13042005.074037.14048/app.query.fasta\_1.1678  
-DB=GenEmbl -QFMT=fastp -SUFFIX=rge -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10051909 @CGN 1 1 9091 @runat.13042005.074037.14048 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	2824	6	AR208565 Sequence
2	3395	83.3	2934	8	AK120560 Oryza sat
3	3155	77.4	2665	8	AY165599 Saccharum
4	3097	76.0	2378	8	AK099716 Oryza sat

5	3052.5	74.9	136267	8	AP005756	AP005756 Oryza sat
6	2825.5	69.3	2800	8	AK102640	Oryza sat
7	2808	68.9	2614	8	HVU534445	AJ534445 Hordeum v
8	2674	65.6	2601	6	AR208568	Sequence
9	2657	65.2	2570	8	AY094465	AY094465 Arabidops
10	2584	63.4	2190	6	AX506620	Sequence
11	2584	63.4	2190	8	ATH532570	AJ532570 Arabidops
12	2583	63.4	2168	8	AK065191	AK065191 Oryza sat
13	2514	61.7	142114	8	AC073166	AC073166 Oryza sat
14	2514	61.7	300957	8	AE017116	AE017116 Oryza sat
15	2502	61.4	86710	8	ATF23E12	AL161587 Arabidops
16	2502	61.4	197859	8	ATCHRIV83	AL161587 Arabidops
17	2368	58.1	135583	8	AC136843	AC136843 Oryza sat
18	2275	55.8	2205	6	AX412656	Sequence
19	2275	55.8	2205	6	AX507559	Sequence
20	2245	55.1	2190	8	ATH532571	AJ532571 Arabidops
21	2217	54.4	2426	8	ATUGTRPR	250752 A.thaliana
22	2191.5	53.8	2516	8	HVU534446	AJ534446 Hordeum v
23	2144	52.6	103192	8	AC007369	AC007369 Arabidops
24	2144	52.6	143879	8	AC069251	AC069251 Genomic o
25	2128	52.2	114918	8	AP004945	AP004945 Lotus cor
26	2117	52.0	94349	8	ATF26013	AL133452 Arabidops
27	2093.5	51.4	100987	2	AC121239	Medicago
28	1968	48.3	120743	8	AP004082	Oryza sat
29	1944.5	47.7	115134	8	AC126786	Medicago
30	1918	47.1	121112	8	AC144482	Medicago
31	1670.5	41.0	1692	6	AR208569	Sequence
32	1557.5	38.2	87592	8	AC099739	Oryza sat
33	1557.5	38.2	127098	8	AC144426	Oryza sat
34	1557.5	38.2	154128	8	AP000615	Oryza sat
35	1470.5	36.1	1564	8	AK067391	Oryza sat
36	1436	35.2	2661	8	AK120270	Oryza sat
37	1424	34.9	1487	6	AR208571	Sequence
38	1411	34.6	122809	8	AC127019	Medicago
39	1409	34.6	2516	8	AK068224	Oryza sat
40	1406	34.5	1625	8	AK059705	Oryza sat
41	1174	28.8	1315	8	BT009593	Triticum
42	1159.5	28.5	70311	8	AF128457	Oryza sat
43	1140	28.0	77605	8	AF119222	Oryza sat
44	1140	28.0	142852	8	AF161269	AF161269 Oryza sat
45	973	23.9	106246	8	AF528565	AF528565 Zea mays

ALIGNMENTS

RESULT 1  
LOCUS AR208565 2824 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 1 from patent US 6383776.  
ACCESSION AR208565  
VERSION AR208565.1 GI:21509752  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2824)  
AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.  
TITLE Plant sugar transport proteins  
JOURNAL Patent: US 6383776-A 1 07-MAY-2002;  
FEATURES  
source  
1..2824  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:	3.94e-225	Length:	2824
Pred. No.:	3517.00	Matches:	684
Percent Similarity:	95.18%	Conservative:	27
Best Local Similarity:	91.57%	Mismatches:	34
Query Match:	86.31%	Indels:	2
DB:	6	Gaps:	2

US-10-051-909-32 (1-800) x AR208565 (1-2824)

QY 56 MetSerGlyAlaValLeuValAlaValAlaSerIleGlyValnLeuLeuGlnGlyTyr 75  
 Db 238 ATGGGGGGCGCGTGTATGCTGCCATCGCGGCTCTATCGGCAACTTGTCTGAGGGCTGG 297  
 QY 76 AspAsnAlaThrIleAlaValLeuTyrIleValLysGluPheGlnLeuGlnAsn 95  
 Db 298 GACATGCCACCAATGCTGGAGCGCTCTGTATACATTAAGAGGAATTCACCTCGAGAGC 357  
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
 Db 358 GAGCCTCTCATGAGGCGCTCATGCGCCATGTTCTCATTTGGGGCAACAGTCAACACA 417  
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
 Db 418 ACATCTCCGGGGCCCAAGGCTGACTCGGTGTAGAGGCCCATCGTGGCTCGGCT 477  
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeuLeu 155  
 Db 478 GTCTCTACTTCTGTCAGTGGCTGGTGTATCGGTGGGGGCCCAATTTGTGTACATCTTGTCT 537  
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
 Db 538 CTGCAAGGCTCATTTGATGGTGGTATCGGTGGGGGCCCAATTTGTGTACATCTTGTCTCTAC 597  
 QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194  
 Db 598 ATCTCGAANAATGCACCGCACAGANATCTTGGGGCTGNTNGAACACCGTTCGCGCAGTTC 657  
 QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213  
 Db 658 ATGGGGCTCAGNGAGGAGTTCCTCTCTACTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGT 717  
 QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe 233  
 Db 718 CCCAAACCTGATTGGAGGCTCATGCTGTGAGTTCGTGATCCCGTCACTTATNTATCTTT 777  
 QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIleGlyArgMet 253  
 Db 778 GGACTGACTGTCTTCTTACTTGGCTGAATCACCAGGTGGCTTGTNAGCAAGAGGAGATG 837  
 QY 254 AlaGluAlaLysValLeuGlnLysLeuValGlyLysAspValSerGlyGluLeu 273  
 Db 838 GCGGAGCGCAAGAGAGTGTGNTGCAAGGCTGCGGGGAAGAGAGATGTCTCANGGGGAGANG 897  
 QY 274 SerLeuLeuLeuGluGlyValGlyGlyAspThrSerIleGluGlyTyrIleIle 293  
 Db 898 GCTCTTCTAGTTGAAGGTTTGGGGTTCGGTAAAGATACACGATTTTNNAGAGTACATCAT 957  
 QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyValAspGlyGluGlnIleThr 313  
 Db 958 GGACCTGCCACCGAGGAGCGGATGATCTTGTAACTGACGCTGATAGGACAAATCAACA 1017  
 QY 314 LeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerIleGlyProIleMet 333  
 Db 1018 CTTTATGGGCGCTGAAGAAGCCAGTCAATGATGCTCGACCTTCAAGGAGCCCATCATG 1077  
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353  
 Db 1078 CTTTGAAGTGTGCTTCTCTTGCATCTCGTATGCGTGGGAGCATGGTGAACACAGAGTGTACC 1137  
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373  
 Db 1138 CTTTATGGATCCGATTTGTGACACTTTTGTGTAGTGTCCATGAGAAATATGCTCAAGCTGGA 1197  
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
 Db 1198 GGAAGTATGAGGAGCACATTTTCCAACTTTGGAAGTATGTTTCACTGTGTACAGATCAG 1257  
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGlyTyrAla 413  
 Db 1258 CATGCCAAATGAGCAGTGGGATGAAGAGATCTTCATAGGAGTACAGGAGTACGCA 1317

QY 414 SerAspGlyAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433  
 Db 1318 TCTGTATGGTGCAGAGAGTGTATGAGGACAAATCTCATAGCCCATTTGTGTCCAGGCAG 1377  
 QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453  
 Db 1378 GCAACAGGTGCGGAAGGAGGACATTTGTGCACCATGTGTCCAGGAGTGTCTTTGAGC 1437  
 QY 454 MetArgArgGlnThrLeuLeuGlyGlyAspGlyValSerSerThrAspIleGly 473  
 Db 1438 ATGAGAAGGCAAGCTCTTAGGGAGGCTGGAGATGTGTGAGCAGCACTCATATCGGT 1497  
 QY 474 GlyGlyTyrPheGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493  
 Db 1498 GGGGATGGCAGCTTCTTGGAAATGTCTCAGAGAAGAGGTGAGATGTGTAGAAAGGAA 1557  
 QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySer 513  
 Db 1558 GGTGGTTCACAAAGAGTCTATTTGCCAACAGAGGGAGTTCCTGGCTCAAGAAGGGCTCA 1617  
 QY 514 IleValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533  
 Db 1618 ATTTGTTTCACTTCCCGTGGTGGCGATGTTCTGTAGGGGTAGTGAGTTTGTATGCTGTCT 1677  
 QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
 Db 1678 GCTTTAGTAACTCAGTCAGCACCTTTTCTCAAAAGGGTCTTTGCTGAACACGCAATGTCAGAT 1737  
 QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573  
 Db 1738 GCTGCCATGTTTCAACCATCTGAGGTAGTGTGCCAAAGGTTTCAAGTTGGAAGATTTGTTT 1797  
 QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593  
 Db 1798 GAACCTGGAGTGGGGTGGCTGCTTGTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1857  
 QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGluAlaValAla 613  
 Db 1858 GCTGGAAATAACGGTGTCTGTACTATACCCCAAAATTTCTTGAGCAAGCTGGTGGGCA 1917  
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuLeuSerSerLeu 633  
 Db 1918 GTTATTTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCTGTATGATCTCTCTC 1977  
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuMetAspLeuSerGly 653  
 Db 1978 ACTACCTTACTAATGCTTCTTGTGATGGCTTTGGCCATGCTGTATGATCTTTCCGGA 2037  
 QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
 Db 2038 AGAAGGTTTTTGTGCTAGGCACAAATTCCAATCTTGTATGACATCTCTAGTTATCTCTGTT 2097  
 QY 674 ValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
 Db 2098 GTGTCCAAATCTAATTTGATTTGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157  
 QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713  
 Db 2158 ATGCTCTACTTCTGCTGCTTGTGATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217  
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733  
 Db 2218 GAGATCTTTCACACCGAGGTTCTGTGGCTCTGTATGCAATTTGTGCTTCACTCTG 2277  
 QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753  
 Db 2278 ATCGGAGATATCATCTGTCACCTACAGCTTCTCTGTATGCTGTGATGCTATGAGCTGGCG 2337  
 QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
 Db 2338 GGTGTTTTTTCAGCATATATGAGTCTGTATGCTTGTATTTCTTGTGTTGCTTCTCTTAAG 2397

QY	774	ValProGluThrIysGlyMetProLeuGluValIleThrGluPheAlaValGlyAla	793
Db	2398	GTCCCTGAGACAAAGGGAGTCCCTTGGATTATTACCGAATCTTTGCGAGTTGGTGGC	2457
QY	794	LysGlnAlaAlaAlaLysAla	800
Db	2458	NAGCAAGCGGTGCAAAAGCC	2478
RESULT 2			
ACCESSION	AK120560	1	GI:37990183
VERSION	FLI CDNA; CAP trapper.		
KEYWORDS	Oryza sativa (japonica cultivar-group)		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1		
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ohtsuka, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kuroaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science	301 (5631),	376-379 (2003)
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Oeato, N., Ota, Y., Ohtsuka, K., Satoh, K., Sasaki, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.		
TITLE	Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice		
JOURNAL	Unpublished		
REFERENCE	3		
AUTHORS	Kikuchi, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JAN-2003)		Shoeshi Kikuchi, National Institute of

Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)			
COMMENT	This clone is one of the 32K full-length cDNA clones from japonica rice.		
URL : http://cdna01.dna.affrc.go.jp/cDNA/NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Yamamoto, M. and Nakahama, Y.			
FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.			
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takahashi, F., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.			
FEATURES	Location/Qualifiers		
source	1..2934		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="mRNA"		
	/cultivar="Nipponbare"		
	/db_xref="taxon:39947"		
	/clone="J013132C18"		
ORIGIN			
Alignment Scores:			
Pred. No.:	5,826-217	Length:	2934
Score:	3395.00	Matches:	676
Percent Similarity:	86.84%	Conservative:	43
Best Local Similarity:	81.64%	Mismatches:	76
Query Match:	83.31%	Indels:	34
DB:	8	Gaps:	5
US-10-051-909-32 (1-800) x AK120560 (1-2934)			
QY	3	SerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgGluArg	22
Db	38	AGTCGTGGGTGGTCTTGGCGGACGCCCGC---GCCCGCGCACCGCTGGCGCGTGGCGGT	94
QY	23	LeuLeuProSerVal-----ValLeuAla	30
Db	95	GGCTGTGTCTACCGCCACGAGAGGATCCGATCGCCCGCGCGGTGTCTCTGCT	154
QY	31	Leu-----ProGlyProLeuProProAlaSerCysSerSerGlnGluPro-----	45
Db	155	CTTCTCTCTCCGCTGGCGCGGACAAAGCGGGATCTTGGCGGCGCTT-CCTCCAGGGAT	213
QY	46	-----ValThrSerAsp-----Aap	50
Db	214	CTCGCTCGCGCGCGCGCTCGTCTGCTCGTCCGCCCAAGAGTGGGAAGAGGGAT	273
QY	51	IleLeuGluAspIysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAen	70
Db	274	CGGAGAGGGGTGAAGATGTCGGGTGCTGCTCTGTGCAATCGCGGCTTCATCGGCAC	333
QY	71	LeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrlleYsLeuGlu	90

Db 334 CTGCTGAGGGATGGCAATGCCACCATTCAGGTGCTACTATACATAAAGAGGAA 393  
Qy 91 PheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGly 110  
Db 394 TTCAGAGCTAGAAGTGAGGCCACTTGGAGGGGCTAATCGTGGCCATGCTCACTGATTGGT 453  
Qy 111 AlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyValArgProMet 130  
Db 454 GCACCATCATCACAACTTTCAGGGCCGGTATCAGACTGATCGGCGCGCCCTATG 513  
Qy 131 LeuIleLeuSerSerIleLeuThrPhePheSerGlyLeuIleMetLeuThrSerProAsn 150  
Db 514 CTCATTCTCTTCAATCTCTACTTCTCTCAGCAGCCCTCATCATGCTATGCTGCCCAAT 573  
Qy 151 ValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170  
Db 574 GTCTATGTGTTACTGCTCGCAGCCCTTATAGATGGATTTCGGCATCGGCTTGCTCACA 633  
Qy 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190  
Db 634 CTTGTACCTTTGTACATCTCAGAGCAGCTCTCTCAGAGATCAGGGGTTTCTGAAATACA 693  
Qy 191 LeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMet 210  
Db 694 CTGCCACAGTTCAGTGGGTTCAGAGGGATGTTCTTGTCTCTACTGCTGCTGTTGGGATG 753  
Qy 211 SerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeu 230  
Db 754 TCACGTGTCCTGCTACCTGACTGGAGAAATATGCTTGGTGTCTCGCAATCCCTTCGTTG 813  
Qy 231 PhePheGlyLeuThrIlePheTyrIleuProGluSerProArgTrpLeuValSerIys 250  
Db 814 TTTTTCCTCGGATTCACAATATCTATCTGCCAGAAATCACCAAGATGGCTTGTCAACAA 873  
Qy 251 GlyArgMetAlaGluAlaIysIysValLeuGlnIysLeuArgGlyLysAspValSer 270  
Db 874 GGGCGGATGGCTGAGCAAAAGAGGTATTACAAAATTTACGTGGAGAGAGGATGTCTCA 933  
Qy 271 GlyGluLeuSerLeuLeuLeuGlyLeuGluValGlyGlyAspThrSerIleGluGlu 290  
Db 934 GGAGAAATGGCTCTTCTTGTGAAGGTTTGGAGGTGGAGCTGACACTTCCATTGAAGAG 993  
Qy 291 TyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIysGlu 310  
Db 994 TACATCATCGGACCGGCTATAGAGCCAGCTGATGAGCATGTTGTGTGATGGCGATAAGGAC 1053  
Qy 311 GlnIleThrLeuTyrGlyProGluGlyGlnSerTrpIleAlaArgProSerIysGly 330  
Db 1054 CAATTAACACTGTACGGGCTGTAGAGGGCCCAATCATGGATTGCTCGACCTTCCAAGGGA 1113  
Qy 331 ProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGln 350  
Db 1114 CCAGCATTTCTTGGAAAGTGTGCTTCTCTTACATCTCGTCATGGCAGCATGGTGNACCAG 1173  
Qy 351 SerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPro 370  
Db 1174 AGCGTGCCACTTATGGATCCTATAGTCAGCTTTTTCGGCAGGTGTCATGAGAACATGCT 1233  
Qy 371 GlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerVal 390  
Db 1234 CATGCTGGAGGAAGCATGCGGAGTACATTTGTTTCCCAACTTGGTAGTATGTTAGTGTG 1293  
Qy 391 ThrAspGlnHisAlaIysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGlu 410  
Db 1294 ACAGATCAGACCCCAAGGTTGATCAATGGGATGAGGAGAACCTTCATAGGGATGATGAG 1353  
Qy 411 GluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeu 430  
Db 1354 GAATACGGCTCTGATGCTGGAGGTGACTATGAAGACAATGTCCACAGCCCGTTGCTG 1413  
Qy 431 SerArgGlnAlaThrClyAlaGluGlyLysAspIleValHisGlyHisArgGlySer 450

Db 1414 TCCCGCAGACACCAGCGCAGAGGAGGAGGCATTTGCACACCATGTCTCATCTGTTGAAAGT 1473  
Qy 451 AlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyAspGlyValSerSerThr 470  
Db 1474 GCTCTGAGTATGAGAAGAAGACCTCTTTGGAAGAGGGTGGCGAGCGGTGAGCAGCACT 1533  
Qy 471 AspIleGlyGlyGlyTyrGlnLeuAlaTrpIysTrpSerGluLysGlyGlyGlyValAsnGly 490  
Db 1534 GGCATTTGGTGGGGATGGCAGCTTGCATGTGAATGGTCAGAGCGAGAGGCGAGATGGT 1593  
Qy 491 ArgLysGlyGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArg 510  
Db 1594 AAGAAGAAAGAGGATTTAAAGGATCTATTGCACCAAGAGGAAGTCCAGGATCAAGA 1653  
Qy 511 ArgGlySerIleValSerLeuProGlyGlyAspValPheGluGlySerGluPheVal 530  
Db 1654 AGGGGCTCAGTATTATTTCACTTCTGTTGGAGGGAGCGCTCTGGAAGGCGAGCAATTCATA 1713  
Qy 531 HisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArg 550  
Db 1714 CATGCTGCCGCTTGGTAAGCCACCCAGCACCTTTACTCCAGGATATCATTTGAACAGCT 1773  
Qy 551 MetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLys 570  
Db 1774 ATGTCCGGTCCAGCCATGATTTCATCCATCAGAGGCGAGCTGCCAAAAGGTTCAAGCTGGAAA 1833  
Qy 571 AspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeu 590  
Db 1834 GATTTGTTTGAACCTGGAGTGAGGGGTGCTTGTATTAGTTCGGTGTTCGAAATTCAAATCCTT 1893  
Qy 591 GlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAla 610  
Db 1894 CAACAGTTTCTGCTGGAATAAATGGGGTTCCTATTACACTCCACAAATACTTTGAGCAAGCG 1953  
Qy 611 GlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIle 630  
Db 1954 GGGGTGGCAGTCTCTCTTCCATCTTGGCCCTCAGTTCCAGCATCAGCTCCATCTTGATC 2013  
Qy 631 SerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuMetAsp 650  
Db 2014 AGTCTCTGACCAACCTACTGATGCTCTAGCATGTCTTACATTTGTTAGCTATGACTTATGAC 2073  
Qy 651 LeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670  
Db 2074 ATCTCTGGAAGAGGTTTCTGCTTCTGGGCAAAATTCAGTCTTGTATGACATCTCTAGTT 2133  
Qy 671 IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690  
Db 2134 GTCTTGGTGTGTCTCAATGTTATTCGACCTGGGTACAGTGGCCCGCCGCGCATCTCCACA 2193  
Qy 691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710  
Db 2194 ATCAGCGTCAATCATCTACTTCTGCTGCTTCGTTCATGGGATTCGGTCCGATCCCAACATT 2253  
Qy 711 LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe 730  
Db 2254 CTGTGTGGGAGATCTTCCCACTAGGGTCCGGGCGCATCTGCATTCGCATCTCGCCCTG 2313  
Qy 731 ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750  
Db 2314 ACATTCGTGAATGGTGAATATCATGCTACCTACAGCTTCCCTGTGATGCTGATGCCATC 2373  
Qy 751 GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuLeuSerPheValPheVal 770  
Db 2374 GGCTTAGCAGGCTCTTGGTATATACGAGGTGTTTGTCTCGATTCCTTCGTGTTCGTC 2433  
Qy 771 PheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAla 790  
Db 2434 TTCTCAAGGTCCCGGAGACGAGGCGCATGCCGCTCGAAGTCATCCAGGATCTTTCGCT 2493  
Qy 791 ValGlyValAlaLysGlnAlaAla 798  
Db 2494 GTTGGTGCAAGCAAAATGCGAGCT 2517



RESULT 3	AY165599	2665 bp	mRNA	linear	PLN 07-JUN-2003
LOCUS	Saccharum hybrid cultivar	putative sugar transporter type 2a	mRNA,		
DEFINITION	complete cds.				
ACCESSION	AY165599				
VERSION	AY165599.1	GI:31505503			
KEYWORDS	Saccharum hybrid cultivar				
SOURCE	Saccharum hybrid cultivar				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.				
REFERENCE	1 (bases 1 to 2665)				
AUTHORS	Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.				
TITLE	Identification of a novel sugar transporter homologue strongly expressed in maturing stem vascular tissues of sugarcane by expressed sequence tag and microarray analysis				
JOURNAL	Plant Mol. Biol. 52, 371-386 (2003)				
REFERENCE	2 (bases 1 to 2665)				
AUTHORS	Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-OCT-2002) Division of Plant Industry, Commonwealth Scientific and Industrial Organisation, 120 Meiers Rd, Indooroopilly, QLD 4068, Australia				
FEATURES	Location/Qualifiers				
source	1..2665				
	/organism="Saccharum hybrid cultivar"				
	/mol_type="mRNA"				
	/db_xref="taxon:120810"				
	97..2334				
	/codon_start=1				
	/product="putative sugar transporter type 2a"				
	/protein_id="AAO37640.1"				
	/db_xref="GI:31505504"				
	/translation="MSGAAVAIVASIGNLQGWDNATISAALIYIKKFKLESEPTV ERLIVAMSLGATVLTITSGPVDWIGRLPMLITLSSVLYFVSSLIWSPNVYVLLIA RLINGGVGLAVTLVLYISETAPPEIRGLNLTLPQFSGSGMFLSCVMFGMSLLPS POWRIMGLVLPSPFLFFGLTIFYPESPRMLVSKGRMAEKVQLRLRGKDDVTGEM ALLIVEGVGGTFTFEITYTGAIRPADHDVADGRERIILFPGEEGQSWTARPSKGS SMLGSVLGSGHSGWVQNLPLTDPVITLPGSVHENISLAGGSWBRITVPSKFDMSLS IDDPKTKQWDEENLHRNEYATDGAASDYNNVHPSLLSRQNTSAREGIAHHGH RSGLGLRQSDSDEGEAATSTGIGGWQALWRSEREGDCKKEGSKFKRIYLHQSG VAGRSRVSVLPFGQSDASGEGKFIHAALVQSALYPRDITEORMAGPATMHPSEAA AKVPSWRDLFEFGVRALLVGGIQTILQAFAGINGVLYTPIQIMEQAGVAVLISNLGL SASASILISSTVAILMLPSIGLAMRLMDVSRRLLSLTIPLVLIASLVLVWSNVIE LGTVHAVLSTVLTILVLCFKWGGPPIILICAEFFFTVRVGGICIALICALIFWVGDI ITVYSVFWMLNAILGEGVFGIYAVACAIAPFVYLKVPETKGMPLVETEFVAVGAKQ AVAKA"				
	97..156				
misc_feature	/notes="putative; transmembrane-region site"				
misc_feature	115..774				
	/notes="pfam00083; Region: sugar transporter"				
misc_feature	238..297				
	/notes="putative; transmembrane-region site"				
misc_feature	313..372				
	/notes="putative; transmembrane-region site"				
misc_feature	406..462				
	/notes="putative; transmembrane-region site"				
misc_feature	511..570				
	/notes="putative; transmembrane-region site"				
misc_feature	589..654				
	/notes="putative; transmembrane-region site"				
misc_feature	658..1794				
	/notes="Region: large cytoplasmic loop"				
misc_feature	1624..2298				
	/notes="pfam00083; Region: sugar transporter"				
misc_feature	1672..1731				
	/notes="putative; transmembrane-region site"				





QY 523 ValPheGluClySerGluPheValHisAlaAaLeuValSerGlnSerAlaLeuPhe 542  
 Db 1201 GTCCTGAAGGACGCAATTCATACATGCTGCGCTTTGGTAAGCCACGACCTTTAC 1260  
 QY 543 SerIlyGlyLeuAlaGluProArgMetSerAspAlaAaMetValHisProSerGluVal 562  
 Db 1261 TCCAAAGATATATTAAGACAGCGTATGTCCGGTCCAGCCCATGATTCATACATCAGAGCA 1320  
 QY 563 AlaAlaIlyGlySerArgTrpIlyAspLeuPheGluProGlyValArgAlaLeuLeu 582  
 Db 1321 GCTGCCAAAGGTTCAAGCTGGAAGATTGTGTTGAACCTCGAGTGGCGGTGGTTGTTA 1380  
 QY 583 ValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyr 602  
 Db 1381 GTCGGTGTGGAAATTCATCAACAGTTGCTGGAATAAATGGGTTCTCTATTAC 1440  
 QY 603 ThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerIlyPheGlyLeuSer 622  
 Db 1441 ACTCCCAATATCTTGAGCAGCGGGGTGGAGTTCCTTCCAACTTGGGCTCAGT 1500  
 QY 623 SerAlaSerAlaSerIleIleSerSerIleThrLeuLeuMetLeuProCysIle 642  
 Db 1501 TCGACATCAGCTTCATCTGTATCAGTCTCTGACACCCCTACTGATGCTTCCTAGCATT 1560  
 QY 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgPheLeuLeuGlyThrIle 662  
 Db 1561 GGTTCAGCCATGAGACTTATGACATCTCTGGAAGAGGTTTCTGCTTCGGGACCAATT 1620  
 QY 663 ProIleLeuAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThr 682  
 Db 1621 CCACTCTTGATAGCATCTCTAGTCTGCTGTGTTGTCCTCAATGTTATCGACTGGGTACA 1680  
 QY 683 LeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMet 702  
 Db 1681 GTGGCCACGCGCACTCTCCCAATCAGCGGTCACTACTTCTGCTTCTGTCATG 1740  
 QY 703 GlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGly 722  
 Db 1741 GGATTCGGTCCGATCCCAACATCTGTGTGCGGAGATCTTCCCACTAGGTCTCGGCGC 1800  
 QY 723 LeuCysIleAlaIleCysAlaPheThrPheThrIleGlyAspIleValThrTyrSer 742  
 Db 1801 ATCTGATTCGCACTGTGCGCTCGACATCTGGATGTGGATATCATATGATCCTACAGC 1860  
 QY 743 LeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValVal 762  
 Db 1861 CTCCTGTGATGCTGAATGCCATCGGCTAGCAGTGTCCTTGGTATATAGCGAGTTGTT 1920  
 QY 763 CysLeuIleSerPheValPheValPheLeuIlyValProGluThrIlySerGlyMetProLeu 782  
 Db 1921 TGCTCGATTGCTTCGTGTTCTGCTCTCTCAAGGTCCCGAGAGCAAGGCGGATGCCGCTC 1980  
 QY 783 GluValIleThrGluPheAlaValGlyAlaIlyValGlyAlaAaAa 798  
 Db 1981 GAAATCATCACCGAGTCTCTGCTGTTGTTGTCAAAGCAAAATGCAGGCT 2028  
 RESULT 5  
 LOCUS AP005756  
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,  
 BAC clone:OSJNBb0035N08.  
 ACCESSION AP005756  
 VERSION AP005756.3 GI:49388934  
 KEYWORDS  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoideae; Oryza.  
 REFERENCE 1  
 Sasaki, T., Matsumoto, T. and Katayose, Y.  
 Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC  
 clone:OSJNBb0035N08

JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

Published Only in Database (2002)  
 2 (bases 1 to 136267)  
 Sasaki, T., Matsumoto, T. and Katayose, Y.  
 Direct Submission  
 Submitted (18-SEP-2002) Takuji Sasaki, National Institute of  
 Agrobiological Sciences, Rice Genome Research Program, Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT

On Jun 28, 2004 this sequence version replaced gi:42627749.  
 Genes were predicted from the integrated results of the following:  
 GENSCAN (http://ccr-081.mit.edu/GENSCAN.html), FGENESH  
 (http://www.softberry.com/), GeneMark\_hmm  
 (http://opal.biology.gatech.edu/GeneMark/), GlimmerM  
 (http://www.tigr.org/db/glimmer/glmr\_form.html), RiceHMM  
 (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor  
 (http://biomformatics.fasstate.edu/cgi-bin/sp.cgi), sim4  
 (http://glabin.cse.psu.edu/html/docs/sim4.html), gap2  
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The  
 genomic sequence was searched against NCBI NonRedundant Protein  
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA  
 sequence database at RGP or DBJ. Protein homologies of the coding  
 regions were searched against NCBI NonRedundant Protein database  
 with BLASTP. ESTs represent the identified cDNA sequences using  
 BLASTN with the corresponding DBJ accession no. and RGP clone ID.  
 Full-length cDNAs represent the identified cDNA sequences using  
 BLASTN with the corresponding DBJ accession no.  
 A gene with identity or significant homology to a protein is  
 classified based on the protein name to indicate the homology level  
 such as same name, 'putative-' and '-like protein'. A gene without  
 significant homology to any protein but with full-length cDNA or  
 EST homology (covering almost the entire length of partial  
 sequence) is classified as an 'unknown' protein. A gene predicted  
 by two or more gene prediction programs is classified as a  
 'hypothetical' protein according to IRGSP standard. A gene  
 predicted by a single gene prediction program is also classified as  
 a probable 'hypothetical' protein and is included as a  
 miscellaneous feature of the sequence.  
 The orientation of the sequence is from M13rev to -21M13 of the BAC  
 clone. This sequence of OSJNBb0035N08 clone has an overlap with  
 P0620H05 (DBJ: AP005394) clone at 5' end and with OJ1711.D06  
 (DBJ: AP004857) clone at 3' end. Detailed information on overlap  
 and assembly quality together with annotation of this entry is  
 available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.  
 Location/Qualifiers  
 1. .136267  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="2"  
 /clone="OSJNBb0035N08"  
 complement(join(157..256,518..608,645..687,1665..1820))  
 /gene="OSJNBb0035N08.1"  
 complement(join(157..256,518..608,645..687,1665..1820))  
 /gene="OSJNBb0035N08.1"  
 /note="hypothetical ORF  
 predicted by GENSCAN  
 this category is not included in IRGSP standard"  
 complement(6895..7272)  
 /gene="OSJNBb0035N08.2"  
 complement(6895..7272)  
 /gene="OSJNBb0035N08.2"  
 /note="hypothetical ORF  
 predicted by GlimmerM  
 this category is not included in IRGSP standard"  
 complement(11805..13042)  
 /gene="OSJNBb0035N08.3"  
 complement(join(<11805..12148,12855..>13042))  
 /gene="OSJNBb0035N08.3"  
 /note="supported by full-length cDNA (s): AK111070"  
 complement(11805..13042)  
 /gene="OSJNBb0035N08.3"

FEATURES  
 source

gene  
 misc\_feature  
 gene  
 misc\_feature  
 gene  
 mRNA  
 misc\_feature



QY 81 ----- 81  
Db 48860 GTGTCGTCTAGATTGGTTATTGGTTGAGAGTTTGTGAGTACTAGGAAATATGAACAGC 48919  
QY 81 ----- 81  
Db 48920 TGTCAATTTTTTCTTAATATCAGTGTGTGGTGACCTCTTAATAATTTTTCAGATTAAATAGA 48979  
QY 81 ----- 81  
Db 48980 CTGCTTTAGATCAGATGTTTGTGTGAGTCTAATTATGTTCTTCTAGTAACATGTCGTGG 49039  
QY 81 ----- 81  
Db 49040 TTGTTGTTTGGTCTTAAGAGCAATTTGGCACAATTTGTTTATCAGCAAACTGGATGCGCATG 49099  
QY 81 ----- 81  
Db 49100 ATATTAAATAGTCTCTTAATTTTTTCTAGATCTGTACTAATCTTATTGAAAGAAAGATGTT 49159  
QY 81 ----- 81  
Db 49160 CTAATCATAAATGTTAGCTATTGTTGTGTGTTGTAATCATTTACTGTTTGGCCATTTCTTTG 49219  
QY 81 ----- 81  
Db 49220 TTCTGAACACCTGTAAAGATTGTTCTATCCAACAAGAAAGTGTAAATGTTTAAACAGGCAC 49279  
QY 82 -----AlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
Db 49280 AATTATTAAACTTCTTCAGGTGCTGTACTATACATAAAGAGGAATTCAGCTAGAAGT 49339  
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Db 49340 GAGCCACATGTGGAGGGCTAATCTGTGCCATGTCATGATGGTGGTCAACCACTCATCA 49399  
QY 116 ThrPheSerGlyProLeuSerHisPheSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
Db 49400 ACATTTCTCAGGCGCGGTATCAGACTGGATTCGGCCCGCCCTATGCTCAITCTCTCTCA 49459  
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeuLeu 155  
Db 49460 AVTCTCTACTCTCAGCAGCCCTCATCATGCTATGTTCCCGAATGCTATGTTGTACTG 49519  
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
Db 49520 CTGCAACGCTTATAGATGGATTTCGGCATCGCTTGGCTGTCAACATTTGTACCTTTGTAC 49579  
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
Db 49580 ATCTCAGAGACAGCTCTCTCAGAGATCAGGGTTTGTGTAATACACTGCCACAGTTCAGT 49639  
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
Db 49640 GGGTCAGAGAGGATGTTCTTGTCTCTGATCGATGGTGTGGGATGTCATGTTGCCGTCA 49699  
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
Db 49700 CCTGACTCAGAAATATGCTTGGTGTTCTCGCAATCCCTTCGTTGTTTCTTCGGATTG 49759  
QY 236 ThrIlePheTyrArgIleMetLeuSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
Db 49760 ACAATATTCTATCTGCCAGAAATACCAAGATGGCTTGTACGAAAGGGCGGATGGCTGAG 49819  
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSer----- 270  
Db 49820 GCAAGAAAGGTATTACAAAAAATTTACGTGGGAGAGAGGATGTCCTCAGGTATGTTTGTGTA 49879  
QY 270 ----- 270  
Db 49880 AAATTTTGATGTTTCATTTTACCACATATTTTCCAGCTGACCAAGTTTAAAAACAATTTCTCCC 49939

QY 270 ----- 270  
Db 49940 AGATTTCTAAACTCATGGTTTAATACTAGAAACATGTGGCATTTGTTAACTGTTAAATTC 49999  
QY 270 ----- 270  
Db 50000 TTGACACGAACCTTAACATATGCAGAAATAGTGTATTTTGAAGAAAAAGGGTAAAAAGAAA 50059  
QY 270 ----- 270  
Db 50060 TCAAAATCTATTAAAGCGCCTTCTATCCGGTGCTCTTTTGAAAAAGCTTATGAATCTTCAAG 50119  
QY 270 ----- 270  
Db 50120 TAAAAATTACTAAGTACACATAACTACTAATTCATTTTCCATAGAAGGCAGAACATGG 50179  
QY 270 ----- 270  
Db 50180 AATGTACAATCGTTATCTCACCGAGTAGGCCAAATAGAGGTGCCCAACTTTTGTGCTGA 50239  
QY 270 ----- 270  
Db 50240 GTCTAATCTGAGAAAAAATAATCAGGTGATCTAGATAGACATCGCCCTCTTTTTTTTTT 50299  
QY 270 ----- 270  
Db 50300 TTTTGGTAAGGTAACCTCTAGAAATTTGGGGGTGGGACCTGCTGTGTAATAATGTGGATGTG 50359  
QY 270 ----- 270  
Db 50360 GGCGGACCCACGGGTCTTATCTACTCTAGTTTAACTAAATGAACTATATTTTGTCTTTTG 50419  
QY 270 ----- 270  
Db 50420 AGATTTTAAAGTGAATTAAGTCCATCTAGTTGACCTAGAGTGGTCAAGAGGCCCCAC 50479  
QY 270 ----- 270  
Db 50480 AGCGCCATGCCTAAACAAGTGTGTTTCAGAAAGTTAGTAATGATTTATCCTTAAGT 50539  
QY 270 ----- 270  
Db 50540 TCATGCTTTCTCTCAAACTAATTACATAATACATGTGTCAACATCAAAATAGGCATATGCT 50599  
QY 270 ----- 270  
Db 50600 GTTAGCAAGTGGTTTCTCTATTTTCTCTCTACACAGACTAAATGGACAGTATTCGGTT 50659  
QY 270 ----- 270  
Db 50660 GGCTGAGAGCATATAATTGCTGTTATTTTAAAGTTTCATCGAACTTCGCTGTAAAGCTGGTTG 50719  
QY 270 ----- 270  
Db 50720 TCCCGAGCTGCATGTAGCTTTCATTAAAGCTGGGCCCTTATATTTCTGTTATTG 50779  
QY 270 ----- 270  
Db 50780 TTTCTGGCACACCATTCATATTTATTTGCTGCTTTTATTCATAGTTACTACTTATTTATG 50839  
QY 270 ----- 270  
Db 50840 CCAATGTTTATACCAATGTCCAAATAAAATCATGTAATCTACGTGAAAAATAAAGTTGC 50899  
QY 271 -----GlyGluLeuSerLeuLeuGlyLeuGly 281  
Db 50900 ATAATCAVTTCAAAATTTATCAAAATTTGCGCAGAGAAATGGCTCTCTCTGTTGAAGGTTTGA 50959  
QY 281 uValGlyGlyAspThrSerIleGlyGluTyrIleIleGlyProAlaThrGluAlaAAs 301  
Db 50960 GGTTCGAGCTGACACTTCCATTTGAAGAGTACATCATCGACCGGCTATAGAGCCAGCTGA 51019  
QY 301 pAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGlyGly 321

Db 51020 TGAGCATGTTGTTGATGGCCATTAAGGACCAATAAACAACCTGTACGGGCTGGAAGGGCCCA 51079  
Qy 321 nserTrpIleAlaAArgProSerLySgLyProIleMetLeuGlySerValLeuSerLeuAl 341  
Db 51080 ATCATGGATTGCTCGACCTTCCAAAGGACCACGACATCTTTGGAAGTGTGCTTCTCTTAC 51139  
Qy 341 aSerArgHisGlySerMetValAenGlnSerValProLeuMetAspProIleValThrIle 361  
Db 51140 ATCTCGTCATGGCAGCATGGTGAACAGAGCGTGCCACCTTATGGATCCTATAGTCACGCT 51199  
Qy 361 uPheGlySerValHisGluAenMetProGlnAlaGlyGlySerMetArgSerThrLeuPh 381  
Db 51200 TTTTCGCAGTGTCCATGAGAACATGCTCATGCTGGAGGAAGCATCGGAGTACATGTT 51259  
Qy 381 eProAenPheGlySerMetPheSerValThrAspGlnHisAlaLyAenGlnTrpAs 401  
Db 51260 TCCCAACTTTGGTAGTATGTTAGTGTGACAGATCAGCACCCCAAGTTGATCAATGGGA 51319  
Qy 401 pGluGluAenLeuHisArgAspAspGluGluTrpAlaSerAspGlyAlaGlyGlyAspTy 421  
Db 51320 TGAGGAGAACCCTTCATAGGATGATGAGGAATACGGGTCTGTATGGTGTGGAGTGACTA 51379  
Qy 421 xGluAspAenLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyAsA 441  
Db 51380 TGAAGACAATGTCACAGCCCGTGTCTGCCACAGACCACGAGCGCAGAGGAAGGA 51439  
Qy 441 pIleValHisGlyHisArgGlySerAlaLeuSerMetArgGlnAlaThrGlyAlaGluGlyAsA 461  
Db 51440 CATTGCACACCATGCTCATCGTGGAGTGTCTGTAGTATGAGAAGAAGACCTCTTGA 51499  
Qy 461 yGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLy 481  
Db 51500 AGAGGGTGGCGCGGTGAGCAGCATGCGCATTTGGTGGGGATGGCAGCTTGCATGGAA 51559  
Qy 481 eTrpSerGluLySgLyGluAenGlyArgLySgLyGlyPheLyAenGlyValTrLe 501  
Db 51560 ATGGTCAGAGCGAGAGGCGAGGATGGTAAGAAGGAAGGAGGATTTAAAGGATCTACTT 51619  
Qy 501 uHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyG 521  
Db 51620 GCACCAAGAGGAATGCCAGGATCAGAGGGGCTCAGTTATTTTCTCTCTGTTGGAGG 51679  
Qy 521 yAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLe 541  
Db 51680 GCAGCGCTCTGAAGGCGAGCAATTCATACATGCTGCGGCTTTGGTAAGCAACAGCACT 51739  
Qy 541 uPheSerLySgLyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerG 561  
Db 51740 TTACTCCAAGGATPATCATTTGAACAGCGTATGTCGGTCCAGCCCATGATTCATCCATCAGA 51799  
Qy 561 uValAlaAlaLySgLySerArgTrpLySgAspLeuPheGluProGlyValArgArgAlaLe 581  
Db 51800 GGCAGCTGCCAAGGTTCAAGCTGGAAAGATTGTTTGAACCTGGAGTGGGGTGGCTT 51859  
Qy 581 uLeuValGlyValGlyIleGlnIleLeuGln----- 591  
Db 51860 GTTAGTCGGTGTGGAAATTCAAATCCTTCAACAGGTAGATGATACCAATTTCAATAAT 51919  
Qy 591 ----- 591  
Db 51920 CTTTTTTGACTGTCCAACATTGACGGTGTTTAAGATTGTTGTPACTATTAAAGAAATTCAT 51979  
Qy 592 -----GlnPheAlaGlyIleAenGlyValLeuTyTrT 603  
Db 51980 AATTCCTTTTATGTTCTTTCTGTCAGTTGCTGGATAATGGGGTCTCTATTACA 52039  
Qy 603 hrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLySgPheGlyLeuSerS 623  
Db 52040 CTCACAAATACTTTGAGCAAGCGGGGTGGCAGTTCTCTCTTTTCCAATCTTGGCCTCAGTT 52099  
Qy 623 erAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleG 643

Db 52100 CAGCATCAGCTCCCATCTTGATCAGTTCTCTGACCACCCCTACTGATGCTTCTTAGCATTTG 52159  
Qy 643 lypPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654  
Db 52160 GTTTAGCCATGAGACTTATGGACATCTCTGGAAGAAGGTACTAACTTGTTTCTCTTGTTT 52219  
Qy 654 ----- 654  
Db 52220 GCATCATCTTAGTACTATACCAATATATACCAAAACAAGAAATGACCCCAAGTTAGTATTAAAT 52279  
Qy 655 -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleA 667  
Db 52280 ACATCATACATTTGTTTCTTTTAGGTTTCTGCTTCTGGGCACAAATTCAGTCTTCATAG 52339  
Qy 667 laSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaL 687  
Db 52340 CATCTCTAGTTCTTGTGTGTTCATATGACCTGGGTACAGTGGCCACGCCG 52399  
Qy 687 euLeuSerThrValSerValIleValTyPheCysCysePheValMetGlyPheGlyProI 707  
Db 52400 CACTCTCCCAATCAGCGTCATCATCTTCTGCTGCTTCGTTCATGGGATTCGGTCCGA 52459  
Qy 707 leProAenIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaI 727  
Db 52460 TCCCCAACATTTGTGTGGGAGATCTTCCACACTAGGGTCCGGCGCATCTGCATTTGCCA 52519  
Qy 727 leCysAlaPheThrPheTrpIleGlyAspIleValThrTyTrSerLeuProValMetL 747  
Db 52520 TCTGGCCCTGACATTCGTGATTGGTGATATCATAGTCACTACAGCTCCCTGTGATGC 52579  
Qy 747 euAsnAlaIleGlyLeuAlaGlyValPheSerIleTyAlaValValCysLeuIleSerP 767  
Db 52580 TGAATGCCATCGCCTAGCAGGTGCTTTTGGTATATACACAGTTGTTTCTCGATTGGCT 52639  
Qy 767 heValPheValPheLeuLySgValProGluThrLySgLyMetProLeuGluValIleThrG 787  
Db 52640 TGTGTTGCTCTTCTCCTCAGGTCCTCCGAGACAGAGGCGATCGCGTCGAAGTCATCACC 52699  
Qy 787 luPhePheAlaValGlyAlaLySgInAlaAlaA 798  
Db 52700 AGTTCTTCTGCTGTGTGTGCAAGCAAAATGCAGGCT 52734  
RESULT 6  
AKI02640  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) cDNA clone:J033100A10, full insert sequence.  
ACCESSION  
AKI02640  
VERSION  
AKI02640.1 GI:32987849  
KEYWORDS  
FLI CDNA; CAP trapper.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1  
The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team.; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kiehimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Hara, A., Harshidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Obato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.







```

Db      1065 TCATGGTAGTATGGTCAGTCAGGTAAGCCCTTGTGGATCCTGTTGTCAACCCCTTTTGG 1124
Qy      363 ySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAs 383
Db      1125 AAGTGTCATGAGAAGATGCTCGAGATATATGGAAGCATGCGAGACACATTGTTTCTTAA 1184
Qy      383 nPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluG 403
Db      1185 CTTTGGCAGCATGTTTGTGTGGCGGAACACAGCAGCAAGCTTAAAGGTGATTTGGGATGCTGA 1244
Qy      403 uAenLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyValAspTyrGluAs 423
Db      1245 GAGT---CAACGGAGGGTGAAGATATATGATCAGACCATTGGTGGGATGACATTGAGA 1301
Qy      423 pAenLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspLe-- 442
Db      1302 TAGCCTCCAAAGCCCACTTATTCTTCGTCAAGCGACAAGCGTGCAGAGGAAGGAGATCGC 1361
Qy      443 -ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyG 462
Db      1362 TGCACCTCATGGCAGTATAATAGGTGCTGTGGGA-----AGAAGTAGTATGATCTCATGCA 1415
Qy      462 uGlyGlyAspGlyValSerSerThrAspGlyGlyGlyGlyTyrGlnLeuAlaTyrLysTr 482
Db      1416 GGGCGGGAGGCAGTAAGCAGCATGGGCATTTGTGGGGATGCGAGTTGGCTTGGGAATG 1475
Qy      482 pSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHi 502
Db      1476 GACTGAGAGAGAGGTGCAGATGCGCAAGAAAGAGGTGGCTTCCAAGTATCTACTTGCA 1535
Qy      502 sGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAs 522
Db      1536 TGAAGAGGGTGTACAGAGTATCGCAGGGCTCTATACTGTCTATGCTT---GGAGGTGA 1592
Qy      522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPh 542
Db      1593 TGTCTCTCGTGGTGGTGGTGTCTCGAGGAGCTGCTCTGTGAGCAACCTGCTCTTTA 1652
Qy      542 eSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVa 562
Db      1653 CTCTAAGGAATTGATGAAGCAACGCTGTGTCGCCCTGTATGTGTGATCATCTCAGGC 1712
Qy      562 lAlaAlaLysGlySerArgTyrLysAspLeuPheGluProGlyValArgArgAlaLeu 582
Db      1713 AGTTGCTAAGAGTCCAAATGGGCAGACTTATTTCGAACCTGGAGTGAAGCATGCTCTGTT 1772
Qy      582 uValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTy 602
Db      1773 TGTGGCATAGGGATCAAAATCTCGCAACAGTTTGTGGCATTAATGGAGTTCTGTACTA 1832
Qy      602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSe 622
Db      1833 CACTCCAAATTTCTTGAGCAAGCTGGTGTGGTGTCTTCTTTCGCAACATGACCTTAG 1892
Qy      622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIl 642
Db      1893 CTCTCATCTGCATCTATTCTTATTAGCGGACTGCACAACCTTGTGACGCTTCCAGCAT 1952
Qy      642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIl 662
Db      1953 TGTATTGCTATGAGGCTCATGATATGTCTGGAAGAAGGTTTCTTCTCTTCCACAAT 2012
Qy      662 eProfileLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682
Db      2013 CCTATCTCTGATAGCATAGCTATCTTGTATCTGGTCAATATCTGATGATGGGAC 2072
Qy      682 rLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMe 702
Db      2073 CATGGTTTCATGCTCACTGCTCCACAGTCAGTGTCTACTTCTGCTTCTTGTTCAT 2132
Qy      702 tGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGl 722
Db      2133 GGGGTTGGGGCCCTATTCCAAACATTTCTCTGTGAGAGATTTTCCCGACACCGTTCTGG 2192

```

```

Qy      722 yLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSe 742
Db      2193 CATCTCATAGCCATCTGTGCCCTAACATTTGGATCGGTGATCATTTGTGCATACAC 2252
Qy      742 rLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValVa 762
Db      2253 CTTCCTCCGATGCTCAAGCCATTCGATCTCGCTGGAGTGTGTTGGATCTACGACATGGT 2312
Qy      762 lCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLe 782
Db      2313 CTGCATACTGGCTTCTCTGTTTCTTCATGAAGGTGCCGAGACAAAGGCGATGCTCT 2372
Qy      782 uGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796
Db      2373 TGAAGTCATCCGAGTTCTTCTGTCGGAGCAAAAGCAGGCC 2415

RESULT 7
LOCUS   HVU534445                2614 bp    mRNA    linear    PLN 14-FEB-2003
DEFINITION   Hordeum vulgare mRNA for hexose transporter (stp1 gene).
ACCESSION   AJ534445
VERSION     AJ534445.1  GI:26986185
KEYWORDS    hexose transporter; stp1 gene.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.

REFERENCE   1
AUTHORS    Weschke,W., Panitz,R., Gubatz,S., Wang,Q., Radchuk,R., Weber,H. and
            Wobus,U.
TITLE       The role of invertases and hexose transporters in controlling sugar
            ratios in maternal and filial tissues of barley caryopses during
            early development
JOURNAL    Plant J 33 (2), 395-411 (2003)
MEDLINE    22424051
PUBMED     12535352
REFERENCE   2 (bases 1 to 2614)
AUTHORS    Radchuk,R.
TITLE       Direct Submission
JOURNAL    Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene
            Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3,
            Gatersleben, D-06466, GERMANY

FEATURES             Location/Qualifiers
     source            1..2614
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="barke"
                     /sub_species="vulgare"
                     /db_xref="taxon:112509"
                     /tissue_type="developing caryopsis"
                     /clone_lib="lambda ZAP2"
                     /dev_stage="0-15 days after flowering"
     gene              1..2614
                     /gene="stp1"
     CDS               83..2314
                     /gene="stp1"
                     /codon_start=1
                     /product="hexose transporter"
                     /protein_id="CAD58958.1"
                     /db_xref="GI:26986186"
                     /db_xref="GOA:Q8GT52"
                     /translation="MSSGAVLVAIAISGNLQGMNATIAGAVLYIKKEFSLQTOPL
                     ISGLVAMSLIGATVITTFSGAVADAVGRPLLIASSVLYFVSGLVMLAPNVPVYVLL
                     AKLIDGFGSLATVLPVLYISTAFDTRGLNLTLPQPSGSGMFLSYCNVFTMSLMP
                     QPDWRMLGLVSPISLMPFALTYFYPESRMLVSKGRMAEAKRVLQRLGRDVSQGE
                     MALLVEGLVGKDTTFEYIIGDDDLADDLAPDQEKLYKGYAEGSVGIARPVXGM
                     GQALGSLGMSRHSVMYSGKSLVDPLVTLPVTFSGVHERKMPYVSGMRSTLPFNPXSG
                     FSAEQQAADWDASHRDDDYASDHGADDDIEDNLNSPLISROATSVGEKIAAPH
                     GSIMGVSSSMQGGDAVDSMGIWGLGGWOLANKWTEREGADGQEGGFQRIYLHEGVIS
                     GDRRGSILSNPGGDIIPPGGEYIQAAALVSPALYSKDLIEQLLAGPAMVHPSEAVAKG

```

TKWALPEPGVKUAFVIGLQILQOPAGINGVLVYTPQILEQAGVGILLNSIGLSS  
 SASLIISALITLLMLPSIGIAMRDMDSRRFLLSTIFVLIALIULVNVLDVGT  
 MVHAAUSTISIVYVFCFVFMGFGPIFNILCAEIFFTSVRGICIAICALTFFMIGDIIVT  
 YTLFVMLNLAIGLAVFGIYAVVCMIAFVFMKVPETKGMPLVEITFEFSVGAQKQKE  
 ATD"

## ORIGIN

Alignment Scores:  
 Pred. No.: 7,65e-178 Length: 2614  
 Score: 2808.00 Matches: 543  
 Percent Similarity: 84.39% Conservative: 84  
 Best Local Similarity: 73.08% Mismatches: 104  
 Query Match: 68.91% Indels: 12  
 DB: 8 Gaps: 6

US-10-051-909-32 (1-800) x HVU534445 (1-2614)

Qy 57 SerGlyValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAsp 76  
 Db 89 TCGGGCCCGCTGTGGTCGCATAGCGGGTCCATCGGAACCTGCTGAGGGGTGGGAT 148  
 Qy 77 AsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGlu 96  
 Db 149 AATCGGACCATTCAGGTCCGTCCTGTATACATAAAGAGAGTTTCAGCTCGAGACCCAG 208  
 Qy 97 ProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThr 116  
 Db 209 CCCTTGATCGAGGGCTCATCGTGGCCATGTCGCTCATCGAGCGACGGTTATCACGACG 268  
 Qy 117 PheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIle 136  
 Db 269 TTCCTCGGGCGGTGGCTGACGCTGTGTGGTGGCGGCCCTGCTCATCGCTCTGTC 328  
 Qy 137 LeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeuLeuLeu 156  
 Db 329 CTCTACTTTGTGAGTGGCTGTGTGCTCTGGGCGCCACGCTATGTGTCTCTTG 388  
 Qy 157 AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIle 176  
 Db 389 GCCAGGCTCATCGAGGGTTCGGTATCGGTTTGGGAGTCACCCCTGTCTGTACAT 448  
 Qy 177 SerGluIleAlaProSerGluIleArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeu 196  
 Db 449 TCGAGACCCCGGACTGACATTAGAGGCTGTGTGAACAGCTGCGCGCAGTTCCAGTGGG 508  
 Qy 197 SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerPro 216  
 Db 509 TCAGGAGGAGATGTTCTTTCTTACTGTCATGGTGTATACCATGTCCCTCATGCGCGACGCT 568  
 Qy 217 AspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThr 236  
 Db 569 GACTGGAGAAATCATGCTTGGGGTTTTGTTCGATCCGCTTATGTATTTTGCATTGACT 628  
 Qy 237 IlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMetAlaGluAla 256  
 Db 629 GTCTTTTATTGCCCGAGTCGCCAGATGGCTTGTGAGCAAGGGAAGATGGCTGAGGCC 688  
 Qy 257 LysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeu 276  
 Db 689 AAGCGAGTATTGCANAGACTCGGGGGAAGGAAGATGCTCAGAGAGATGGCCCTCTT 748  
 Qy 277 LeuGluGlyLeuGluValGlyClyAspThrSerIleGluGluTyrIleIleGlyProAla 296  
 Db 749 GTTGAGAGATTGGGTGTGGAAAGACACATTTGAGGAATACATAATTTGAGACCTGAT 808  
 Qy 297 ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGly 316  
 Db 809 GATGAGCTGGCTGATGACGGGCTGGCTCCA---GATCAAGAGAGATTTGAACCTATGGA 865  
 Qy 317 ProGluGluGlyGlnSerTyrIleAlaArgProSerLys-----GlyProIleMetLeu 334  
 Db 866 GCTGAAGAAGGGGTATCTTGGATCGCCGCTCTTGTAGSGGTGGCGCCAAAGTGCACTT 925

Qy 335 GlySerValLeuSerLeuAlaSerArgHisGlySerMetMetValAsnGlnSerValProLeu 354  
 Db 926 GGAAGCGCTTTTGGCCCTCATGCTCGTCATGGAGATATGTTAGTCAGGGTAAATCTCTT 985  
 Qy 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374  
 Db 986 GTTGACCCCTTGTCACTCTCTTCGGAAGTGTCCATGAGAAAGATGCTGAGTAATGGGG 1045  
 Qy 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrSerGlnHis 394  
 Db 1046 AGCATGGGAGCACCTTGTCTTAACTTTGGCAGCATGTTTAGTGTGGCTGAACAGCAG 1105  
 Qy 395 AlaLysAsnGluGlnTrpAspGluGlnLeuLeuHisArgAspAspGluGluTyrAlaSer 414  
 Db 1106 CAGGCTAAAGCTGACTGGGATGCTGAGAGT---CATAGGGATGATGAAGATTATGCACTG 1162  
 Qy 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434  
 Db 1163 GATCATGGTGGCTGATGACATTGAGGATAACCTCAATAGCCGCTTATTTCTCGTCAAGCG 1222  
 Qy 435 ThrGlyAlaGluGlyLysAspIleVal-----HisHisGlyHisArgGlySerAlaLeu 452  
 Db 1223 ACAAGCGTGGAGGGTAAGGAGATTGCTGCACCTCATGGAAGCATAAATCGGTGGTGAA 1282  
 Qy 453 SerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIle 472  
 Db 1283 AGTAGTAGCATGCG-----GGAGGGAGCGCAGTAAGCAGCATGGGCATT 1327  
 Qy 473 GlyGlyGlyTyrTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLys 492  
 Db 1328 GGTGGGGGTGGCAGCTAGCTTGGAGTGGAGTGGAGTGGAGAGGTGCAGATGGGCAAAA 1387  
 Qy 493 GluGlyGlyPheLysArgValTyrLeuHisGlnGlnGlyValProGlySerArgArgGly 512  
 Db 1388 GAAGTGGGCTCCAGCGATTACTTGTGATGAGAGGGTGTGCAGGTGATCGGAGAGGC 1447  
 Qy 513 SerIleValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAla 532  
 Db 1448 TCTATATTGTCTATGCCA---GGAGGTGATATTCTCTCTGTTGTGATGATATATCCAGGCA 1504  
 Qy 533 AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSer 552  
 Db 1505 GCGCGCTAGTGGAGCAACCTGCTCTTTACTCGAAGGACCTGATAGAGCAGCAGCTTGTCT 1564  
 Qy 553 AspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeu 572  
 Db 1565 GGTCGAGCGCATGGTACATCCATCCGAGCGAGTTCGCAAGGGTACAAAAGTGGCAGAACTA 1624  
 Qy 573 PheGluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnGln 592  
 Db 1625 TTTGAACCTGGAGTGAAGCATGCATGTTTGTTCGCAATTGATTTACAGATCCTCGACAG 1684  
 Qy 593 PheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyVal 612  
 Db 1685 TTTGGGGGTATCAACGGNGTCTCTGTATTACCGCTCAGATACTTGGAGCAAGCAGGTGT 1744  
 Qy 613 AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSer 632  
 Db 1745 GGGATCTTCTATCAACATTGGACTAAGCTCTTCTCTCTCATCTATTTATTTAGTGCC 1804  
 Qy 633 LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer 652  
 Db 1805 TTGCAACCTTGGTGGATGCTTCCAGCATTTGGCATGGCATGCCATGAGACTCATGGATATGTCA 1864  
 Qy 653 GlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672  
 Db 1865 GGAAGAAGGTTTCTCTCTCTCAACATCCCTGTCTTGATAGTGGCAGCTAGCTATCTTG 1924  
 Qy 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692  
 Db 1925 GTTTTGTGGTGAATGTTCTGGATGTGGGAACCATGGTGCAGCCGCTCTCTCAACAGCATCAGT 1984  
 Qy 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712

```

Db      1985  GTCATAGTCTACTTCTGCTTCTTCGTCATCGGGTTTGGACCTATCCCGAATATCTCTGC 2044
Qy      713  AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
Db      2045  GCGGAGATTTCCTCCACCTCCGTCGGTGGCATCTGCATTGCCATCTGGCGCACTAACCTTC 2104
Qy      733  TrpIleGlyAspIleValThrYrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db      2105  TGGATTGGTGACATCATCGTGACATACACTTTCCTCCGTGATGTCACCGCATCGGACTC 2164
Qy      753  AlaGlyValPheSerIleTyAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db      2165  GCTGGAGTTTTCGGAATATATGCTGCTTGTGATGATGACCTTCGTATTCGTCTACATG 2224
Qy      773  LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792
Db      2225  AAGGTCCTTGAGACAAAGGCGCATGCCCTGGAGGTCTATCACCAGGTTCTTCTCGGTCCGG 2284
Qy      793  AlaLysGln 795
Db      2285  GCGAAGCAG 2293

RESULT 8
AR208568
LOCUS      AR208568                2601 bp      DNA      linear      PAT 20-JUN-2002
DEFINITION Sequence 7 from patent US 6383776.
ACCESSION AR208568
VERSION   AR208568.1  GI:21509755
KEYWORDS  Unknown.
SOURCE    Unknown.
ORGANISM  Unclassified.
REFERENCE 1 (bases 1 to 2601)
AUTHORS   Allen, S.M., Hitz, W.D., Kinney, A.J. and Tingey, S.V.
TITLE     Plant sugar transport proteins
JOURNAL   Patent: US 6383776-A 7 07-MAY-2002;
          Location/Qualifiers
FEATURES   Feature             Location
           1. .2601
            /organism="unknown"
            /mol_type="unassigned DNA"

Alignment Scores:
Pred. No.:      6.72e-169      Length:      2601
Score:          2674.00      Matches:      522
Percent Similarity: 82.20%      Conservative: 92
Best Local Similarity: 69.88%      Mismatches:  115
Query Match:    65.62%      Indels:      18
DB:             6              Gaps:         8

US-10-051-909-32 (1-800) x AR208568 (1-2601)
Qy      56  MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTyr 75
Db      175  ATGAAGGTGCCCTCTGTTGCTATTTGGCGCTTCCATTGGTAAATTCCTCCAAAGATGG 234
Qy      76  AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuAsn 95
Db      235  GATAATGCTACCATCCCGGGGTAAATGTTATTAAGNAAGACCTTGTCTTGGGAACA 294
Qy      96  GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db      295  -----ACTATGAAAGGCTTGTGGTGGCATGCTCCCTGATGGACGACGGTAATCACC 348
Qy      116  ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db      349  ACATGCTGCTCATATGCGGATTCGCTCGGTCGGCGACCCCATGATGATGAATCTCATCT 408
Qy      136  IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db      409  GTGCTCTATTCTTGGGTGTTTGGTGATGCTGTGGTCCCAAAATGTGTATGTGTGTGC 468

```

```

Qy      156  LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db      469  TTGGCGAGGCTACTTTGATGGATTGGGATTGGCTTTGCTGTGACTCTTGTCCGGTCTAT 528
Qy      176  IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db      529  ATATCTGAACAGCGCGCTCTGAATAAGGGGTCTGTGAATACGCTTCTCAGTTTCAGT 588
Qy      196  GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db      589  GCGTCTCGAGGAATGTTTGTCTGACTATATGTTTGGCATGTCATTCAGTCCGCG 648
Qy      216  ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db      649  CCTAGCTGAGGCTCATGCTTGGGPTCTGTCTATTCTCTCTCTTGTATTTTGCATTG 708
Qy      236  ThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMetAlaGlu 255
Db      709  ACCATTTTCTTGGCCGAGTCTCTCGTGGTGGTGTGAGCAAGGAAGGATGCTCGAG 768
Qy      256  AlaLysLysValLeuGlnLysLeuArgGlyLysAspPheValSerGlyLeuLeuSerLeu 275
Db      769  GCTAAGAGGTGCTCCAAAGATTGCGCGAAGGAGGATGTGTACGGCGAGATGGCATTG 828
Qy      276  LeuLeuGluGlyLeuValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
Db      829  CTGGTTTGAAGGTCTCGGGATCGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888
Qy      296  AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db      889  GTTGACGATGTGGTGTGATGTCATGNACATGCACAGAGAGAAAGATAAATTCGATTATAT 948
Qy      316  GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db      949  GGATCCCAAGCAGGCGCTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTCTTATTGGC 1008
Qy      336  SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db      1009  -----CTTGGCTGCACACCATGGAAAGCATCATCAACCAAGAGTGCCTCATG 1056
Qy      356  AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db      1057  GATCTCTGTGACACTGTTTGGTAGCATTCATGAGAAGCTCCCCGAGACAGGACAGAGA 1116
Qy      374  GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db      1117  GGAAGCATCGAAGCAGCTCTGTTTCCAAATTTTGGAAAGCATGTTACAGACTGTGAGCCG 1176
Qy      394  HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db      1177  CATGCTAAATTTCAACAAATGGGATGAAGAAAGCTTTACAAAGGGAAACGTGAGGACTACATG 1236
Qy      414  SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db      1237  TCAGATGCAACCGTGGGAGCTCCGATGATATATTTGCACAGTCTTATATCTCACGCCAA 1296
Qy      434  AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeu--- 452
Db      1297  ACAACAAGCCTTGAA---AAAGACTTACCTCTCTCTCTCTCCATGGCAGTATCTCTGGC 1353
Qy      453  SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
Db      1354  AGCATGAGCGCTCACACATAGTCTCATGCAAGGGTCAGGTGAGCAAGGTCGTAGTACAGGT 1413
Qy      472  IleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db      1414  ATTGTGTGTGGTCTGGCACTGGCATGGAATATGACATGATAA---GCTGAGGATGGAANA 1470
Qy      492  LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db      1471  CAACAAGGAGGTTTAAAGGATTTATTTACATGAGGAGGAGTTCCTGCTATCTCTCGT 1530
Qy      512  GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531

```

Db		CGATCCATTGTCATGGTCTCCCGGTGAAGGC-----GAATTTGTCTCAAG	1572
Qy		AlalalaLeuValSerGlnSerAlaLeuPheSerIysGlyLeuAlaGluProArgMet	551
Db		GCTGCTGCTTGTAAGAACAACCGCCTCTTTACTCCAAGGAGCTTATTGATGGACACCCA	1632
Qy		SerAspAlaAlaMetValHisProSerGluValAlaAlaIalsGlysGlySerAgtTrpLysAsp	571
Db		GTTGGCGCTCGAAATGGTTTCACCACATCTGAGACAGCTTCAAAGGGGCCAAAGTTGGAAGAGCT	1692
Qy		LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyVlleGlnIleLeuGln	591
Db		CTTCTTGTAACCAAGGGGTTAAAGCATGTCATTTGTTGTGAGTTGGAATACAAAATACTTTCAG	1752
Qy		GlnPheAlaGlyVlleAsnGlyValLeuTyrrThrProGlnIleLeuGluGlnAlaGly	611
Db		CAGTTTTTCAGGGATAAATGGGGTCTATATATTACACACTCAAAATCCTTTGAAGAGGCCGT	1812
Qy		ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer	631
Db		GTTGAAGTCTCTTCTTCAGATATAGGCATTTGGCTCAGAGTCGGCATCATTTCCATTACGT	1872
Qy		SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuMetAspLeu	651
Db		GCTTTCACAACCTTCTTGATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGGATGTT	1932
Qy		SerGlyArgGpPheLeuLeuGlyThrIleProIleIleuIleAlaSerLeuValIle	671
Db		TCAGGCAGAGGCAGTTGCTACTTACTACAAATCCCCTGCTGATGTGTGCTCACTCATATT	1992
Qy		LeuValValSerAnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal	691
Db		TTGTCAITGGAAACCCCTGGTAATTTGGCAATGTCGCCCATGAGCAATTCCAAAGTA	2052
Qy		SerValIleValTyrrPheCysCysPheValMetGlyPheGlyProIleProAnIleLeu	711
Db		TGCCTTGTGGTTATTATTTCTGCTGTTTGTATGGGTATGACCAATTCCAAACATCCTT	2112
Qy		CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr	731
Db		TGCTCAGAGATATTTCCCCACATAGGTCGTGGCTCTGCACTGCTATCTGTGCATTAGTG	2172
Qy		PheTrpIleGlyAspIleIleValThyrSerLeuProValMetLeuAnlaIleGly	751
Db		TTCTGGATTTCGAGACATCATCACATCTCGCTGCTGTGATGCTCGGCTCTTAGGA	2232
Qy		LeuAlaGlyValPheSerIleTyrrAlaValValCysLeuIleSerPheValPheValPhe	771
Db		CTTGGTGGTGTATTTCGCAATTAACGCAATTTGTTTGTTCATCTCGTGGATATTGTTGTTT	2292
Qy		LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal	791
Db		TTGAAGGTTCCAGAAACAAGGGCATGCCCTTGAAGTCATCTCTGAAATCTTTTCTGTT	2352
Qy		GlyAlaIleGlnAlaAlaLa 798	
Db		GGAGCAAGCAGGCTGCTTCT 2373	
RESULT_9			
AY094465		2570 bp mRNA linear PLN 05-MAY-2002	
LOCUS		Arabidopsis thaliana AT4g53300/F23E12_140 mRNA, complete cds.	
DEFINITION		AY094465	
ACCESSION		AY094465.1 GI:20453188	
VERSION		FLI CDNA.	
KEYWORDS		Arabidopsis thaliana (thale cress)	
SOURCE		Arabidopsis thaliana	
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.	
REFERENCE		1 (bases 1 to 2570)	
AUTHORS		Shinn,P., Chen,H., Cheuk,R.R.Kim,C.J., Meyers,M.C., Banh,J.,	

Percent Similarity:	79.87%	Conservative:	97
Best Local Similarity:	67.52%	Mismatches:	133
Query Match:	65.20%	Indels:	25
DB:	8	Gaps:	10
US-10-051-909-32 (1-800) x AY094465 (1-2570)			
Qy	29	LeuAlaLeuProGlyProLeuProAlaSerCysSerSerGlnGlu-----	44
Db	91	CTTCTTTTCGCT---CCAATTCCAAGTGGATTTTGTTCACCTCACTCATCGCTCGTGTGATC	147
Qy	45	-----ProValThrSerAspAspIleLeuGluAAspLys-----	55
Db	148	AGTGACTCTGTTTTTCGGAATTCCTTCAGATTTCCTTGATATAAAGATAGAACGCGAAT	207
Qy	56	-MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTr	75
Db	208	TATGAGTGGAGCTGTGCTTGTCTATTGCTCTGCTGTGGCACTTGTITACAGGATG	267
Qy	75	pAspAsnAlaThrIleAlaAlaValLeuTyrlleLysLysGluPheGlnLeuGlnAs	95
Db	268	GGATAACGCACTATTTCAGGAGCTGTGTGTACATAAATAAAGAGTTTAATTTGGAGAG	327
Qy	95	nGluProThrValGluGluLeuValSerMetSerLeuIleGlyAlaThrIleValTh	115
Db	328	TAATCCATCAGTGAAGGCTTAATTGTGGCGATGTCACTTATTGTGTCTCTGATTAC	387
Qy	115	rThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSe	135
Db	388	AACATGCTCTGGAGGGTACTGATGGCTGTGTGCGCGTCCCATGCTATATGTCCTC	447
Qy	135	rIleLeuTyrlPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrlValLeuLe	155
Db	448	AATTCCTACTTTTGGTTCTCTAGTAATGCTATGCTTCCGAATGTTTATGTGTGCT	507
Qy	155	uLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTy	175
Db	508	CTTAGGAAGGTTGTTAGATGGATTTGGGGTTGGTCTGTGTGCTCACAACCTTCTCTATTTA	567
Qy	175	rIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSe	195
Db	568	TATATCTGAGACTGCACCACTGAGATTAGGGAGCTGTTGAAATACCGTACCGCAGTTAC	627
Qy	195	rGlySerGlyGlyMetPheLeuSerTyrlCysMetValPheGlyMetSerLeuSerProSe	215
Db	628	TGGCTCTGGAGGATGTTCTTATCTTACTGATGTTTTCGGAATGTCGTTGATGCCATC	687
Qy	215	rProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLe	235
Db	688	ACCTAGCTGGAGATTGATGCTTGTGTCTTTTCATCCCTTCCCTTCTTTTCTCTCT	747
Qy	235	uThrIlePheTyrlLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGl	255
Db	748	CAGGCTCTTCTTCTGCCCGAGTCCCAGAGTGGCTCGTAGCAAAAGGTGCAATGCTTGA	807
Qy	255	uAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyLeuLeuSerLe	275
Db	808	AGCAAGCGGTTCTTCAGAGACTGCGTGTGTCGCGAAGATGTGTCTGCTGAGATGGCTTT	867
Qy	275	uLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrlleIleGlyPr	295
Db	868	GTTGGTTGAGGGTCTTGGAAATTTGGAGGTGAACAACCATAGAGGAATATATAATTGGTCC	927
Qy	295	oAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTy	315
Db	928	CGCGGATGAATGTTACTGATGATCATGATATAGCTGTGGATAGGATCAAAATTAAGTTATA	987
Qy	315	rGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGl	335
Db	988	TGGTGCAGAGAGAGGGCTGAGTTGGTTGCTTAGGCCAGTCAAGGA-----GG	1035
Qy	335	ySerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLe	354
Db		:   :	

1036	AAGCACTATGAGTGTGTTTGTCTCGCCATGGAAGTACAATGACGAGGAGGCAAGGCTCAAT	1095	
Qy	354	uMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGl	374
Db	1096	GATTGATCTCTTGTCACTGTTTGGGAGCGTTTACGAGAAAGATGCCGACACT---GG	1152
Qy	374	ySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHi	394
Db	1153	AAGCATGAGGAGTGCCCTTGTCCACATTTTGGGAGTATGTTTCTGAGTGTGGAGGAATCA	1212
Qy	394	salalyAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrlAlaSe	414
Db	1213	ACCAAGACATGAAGATGGGATGAAGAAATCTTGTGGGAGAGGTGAGGATTATCCATC	1272
Qy	414	rAspGlyAlaGlyGlyAspTyrlGluAspAsnLeuHisSerProLeuLeuSerArgGlnAl	434
Db	1273	CGAC---CATGGAGATGATCTCAAGATGATCTTCACTTCTCCGTTGATCTCAGCTCAAC	1329
Qy	434	aThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSerMe	454
Db	1330	GACAAGCATGGAG---AAAGACATGCCTCACACTGCTCATGGAACTCTTTCTACCTTCAG	1386
Qy	454	tArgAspGlnThrLeuGlyGlyGlyValSerSerValSerSerThrAspIleGlyGl	474
Db	1387	ACATGGAAGTCAAGTSCAGGAGCTCAAGGGGAGGAGCGGTAGTATGCGGATTGGAGG	1446
Qy	474	yGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGl	494
Db	1447	TGGATGGCAAGTGGCATGGAATGGACGGAAGAGAAGATGAATCGGCACAGAAGAAGG	1506
Qy	494	yGlyPheLysArgValTyrlLeuHisGlnGluGlyValProGlySerArgArgGlySerIl	514
Db	1507	TGGGTTTAAACGGATATCTTGCATCAAGAAGGTTTCCCAGGATCTCGACGTGGCTCAAT	1566
Qy	514	eValSerLeuProGlyGlyGlyAspValPheGlyGlySerGluPheValHisAlaAlaAl	534
Db	1567	TGTTTCATTGCTTGGTGGTGTAGGAACCGGTGAG---GCAGATTTTGTACAGCGTCTGC	1623
Qy	534	aLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAl	554
Db	1624	TTTGGTTAGCAACACGCTCTTTATTTCCAAAGACCTTCTCAAGAACAATACAAATGGTCC	1683
Qy	554	aAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGl	574
Db	1684	TGCTATGGTACATCCATCCGAA---ACAACATAAAGGTCATTTTGGCATGATCTTCATGA	1740
Qy	574	uProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAl	594
Db	1741	TCCTGGAGTCAGCGGTGCATTTAGTCGTAGGAGTTGGACTTCAATTAATCTTCAGCAGTTCTC	1800
Qy	594	aGlyIleAsnGlyValLeuTyrlTyrlThrProGlnIleLeuGluGlnAlaGlyValAlaVa	614
Db	1801	AGGCATCAACGGAGTCTTTTACTACACCGCAAAATCCTTGACGAGCGGGTCTCGGAT	1860
Qy	614	lIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuTh	634
Db	1861	CCTACTATCGAAATCGGGATTAGTTCTTCTCCTCAGCATCCTTACTTATAAGTCATTGAC	1920
Qy	634	rThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyAr	654
Db	1921	AACCTTTGTGATGTTTACCTGCAATAGCTGTTGCAATAGAGCTCATGGATCTTTCTGCTCG	1980
Qy	654	gArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVa	674
Db	1981	AAGGACCTTGTCTTCCACCGATACCAATCCTGATAGCATCTCTATTGTTTTAGTAAT	2040
Qy	674	lSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIl	694
Db	2041	CTCAATCTTGTTCATCAACAGCATTTGTGACGCGGTCTTATCAACCGTAAGCGTGT	2100
Qy	694	eValTyrlPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGl	714
Db	2101	GCTCTACTTCTGCTTCTCGTATGGGTTTCGGTCTCTGCTCCAAACATCTCTGTTCAGA	2160



Db	1297	-----GAAGGTTTCCAGGATCTCCAGCGTGGCTCAATT	1329
Qy	515	ValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAa	534
Db	1330	GTITTCATTGGCTGGTGGTGAACCGGTGAG--CGAGATTTTGTACAAGCGTCTGCT	1386
Qy	535	LeuValSerGlnSerAlaLeuPheSerIysGlyLeuAlaGluProArgMetSerAspAla	554
Db	1387	TTGGTTAGCAACACAGCTCTTTATTCAAAGACCTTCTCAAAGNAACATACAATTTGGTCTCT	1446
Qy	555	AlaMetValHisProSerGluValAlaAlaIysGlySerArgTprLysAspLeuPheGlu	574
Db	1447	GCTATGGTACATCCATCCGAA--ACAACCTAAAGGGTCAATTTGGCATGATCTTCATGAT	1503
Qy	575	ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla	594
Db	1504	CCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAAATACCTTCAGCAGTCTCA	1563
Qy	595	GlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal	614
Db	1564	GGCATCAACGGAGTCTTTACTACACACCGAAATCCTTGAGCAGCGGGTGTCCGGATC	1623
Qy	615	IleLeuSerIysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr	634
Db	1624	CTACTATCGAACATGGGGATTAGTTCTTCTCAGCATCCTTACTATTAAAGTCATTGACA	1683
Qy	635	ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg	654
Db	1684	ACCTTTGTGATGTTTACCTGCAATAGCTGTTGCAATGAGGTCATGGATCTTTCTGGTCTGA	1743
Qy	655	ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal	674
Db	1744	AGGACCTTGGCTTCTCACACGATACCAATCCTGATAGCATCTCTATTGGTTTTAGTAATC	1803
Qy	675	SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle	694
Db	1804	TCAAAATCTTGTTCACATGAACAGCATTTGTGCACGCGGCTTATCAACCGTAAGCGTTGTG	1863
Qy	695	ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu	714
Db	1864	CTCTACTTCTGCTTCTTCGTGATGGGTTTCGGTCTCTGCTCCAAACATCCTCTGTTCAGAG	1923
Qy	715	IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle	734
Db	1924	ATTTTTCCAATCGAGTCGCGGGAATCTGCATCCCATCTCGGCATCTACCTCTCTGATC	1983
Qy	735	GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly	754
Db	1984	TGTGACATAATCGTCACTTACAGTCTCCCGTGTGCTCAAAATCCATTGGCAGTAGCTGGT	2043
Qy	755	ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuIysVal	774
Db	2044	GTGTTTCGAATGTACGCAATCGTATGTGCAATTTTCATGGGTCTTTGTTCATTAAAGTC	2103
Qy	775	ProGluThrIysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaIys	794
Db	2104	CCGGAAACTAAAGCATGCCACTTGAAGTCATCACAGAGTCTTTTCTGTGGAGCTAGA	2163
Qy	795	GlnAlaAlaAla	798
Db	2164	CAAGCTGAAGCT	2175

RESULT 11				
ATH532570				
LOCUS	ATH532570	2190 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana mRNA for monosaccharide sensing protein 2			
	(mssp2 gene).			
ACCESSION	AJ532570			
VERSION	AJ532570.1	GI:26800694		
KEYWORDS	monosaccharide sensing protein 2; mssp2 gene.			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			

Eukaryota; Viridiplantae, Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; eurosid II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Stamme, C., Tjaden, J., Trentmann, O., Emmerlich, V. and Neuhaus, E.  
A novel family of plant monosaccharide transporters is involved in a new type of eukaryotic sugar sensing  
Unpublished  
2 (bases 1 to 2190)  
Neuhaus, E.  
Direct Submission  
Submitted (04-DEC-2002) Neuhaus E., Plant Physiology, University of Kaiserslautern, Erwin-Schroedinger-Str. 22, 67663 Kaiserslautern, GERMANY

## FEATURES

gene

**CDS**

## ORIGIN

Alignment Scores:					
Pred. No.:	5,349-163	Length:	2190		
Score:	2584.00	Matches:	512		
Percent Similarity:	80.78%	Conservative:	123		
Best Local Similarity:	68.84%	Mismatches:	893		
Query Match:	63.41%	Indels:	20		
DB:	8	Gaps:	8		

IIS-10-051-909-32 (1-800) x ATH532570 (1-2190)

Or 56 MetSerG|vAlaVal|LeuValAla|LeValAlaSerI|eG|vAsn|Leu|LeuG|nG|vTrp 75

1 ATGAGTGGAGCGTGTGCGTTGTGCTATTGGCTGCTGTGGCAACTTGTTCACAGGATGG 60

Q. And you're not going to tell me that the defendant was not the person who was in the car, is that right?

[illegible][illegible][illegible]

121

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

1. **Introduction**  
 2. **Background**  
 3. **Methodology**  
 4. **Results**  
 5. **Discussion**  
 6. **Conclusion**  
 7. **References**  
 8. **Appendix**  
 9. **Figure 1**  
 10. **Figure 2**  
 11. **Figure 3**  
 12. **Figure 4**  
 13. **Figure 5**  
 14. **Figure 6**  
 15. **Figure 7**  
 16. **Figure 8**  
 17. **Figure 9**  
 18. **Figure 10**  
 19. **Figure 11**  
 20. **Figure 12**  
 21. **Figure 13**  
 22. **Figure 14**  
 23. **Figure 15**  
 24. **Figure 16**  
 25. **Figure 17**  
 26. **Figure 18**  
 27. **Figure 19**  
 28. **Figure 20**  
 29. **Figure 21**  
 30. **Figure 22**  
 31. **Figure 23**  
 32. **Figure 24**  
 33. **Figure 25**  
 34. **Figure 26**  
 35. **Figure 27**  
 36. **Figure 28**  
 37. **Figure 29**  
 38. **Figure 30**  
 39. **Figure 31**  
 40. **Figure 32**  
 41. **Figure 33**  
 42. **Figure 34**  
 43. **Figure 35**  
 44. **Figure 36**  
 45. **Figure 37**  
 46. **Figure 38**  
 47. **Figure 39**  
 48. **Figure 40**  
 49. **Figure 41**  
 50. **Figure 42**  
 51. **Figure 43**  
 52. **Figure 44**  
 53. **Figure 45**  
 54. **Figure 46**  
 55. **Figure 47**  
 56. **Figure 48**  
 57. **Figure 49**  
 58. **Figure 50**  
 59. **Figure 51**  
 60. **Figure 52**  
 61. **Figure 53**  
 62. **Figure 54**  
 63. **Figure 55**  
 64. **Figure 56**  
 65. **Figure 57**  
 66. **Figure 58**  
 67. **Figure 59**  
 68. **Figure 60**  
 69. **Figure 61**  
 70. **Figure 62**  
 71. **Figure 63**  
 72. **Figure 64**  
 73. **Figure 65**  
 74. **Figure 66**  
 75. **Figure 67**  
 76. **Figure 68**  
 77. **Figure 69**  
 78. **Figure 70**  
 79. **Figure 71**  
 80. **Figure 72**  
 81. **Figure 73**  
 82. **Figure 74**  
 83. **Figure 75**  
 84. **Figure 76**  
 85. **Figure 77**  
 86. **Figure 78**  
 87. **Figure 79**  
 88. **Figure 80**  
 89. **Figure 81**  
 90. **Figure 82**  
 91. **Figure 83**  
 92. **Figure 84**  
 93. **Figure 85**  
 94. **Figure 86**  
 95. **Figure 87**  
 96. **Figure 88**  
 97. **Figure 89**  
 98. **Figure 90**  
 99. **Figure 91**  
 100. **Figure 92**  
 101. **Figure 93**  
 102. **Figure 94**  
 103. **Figure 95**  
 104. **Figure 96**  
 105. **Figure 97**  
 106. **Figure 98**  
 107. **Figure 99**  
 108. **Figure 100**  
 109. **Figure 101**  
 110. **Figure 102**  
 111. **Figure 103**  
 112. **Figure 104**  
 113. **Figure 105**  
 114. **Figure 106**  
 115. **Figure 107**  
 116. **Figure 108**  
 117. **Figure 109**  
 118. **Figure 110**  
 119. **Figure 111**  
 120. **Figure 112**  
 121. **Figure 113**  
 122. **Figure 114**  
 123. **Figure 115**  
 124. **Figure 116**  
 125. **Figure 117**  
 126. **Figure 118**  
 127. **Figure 119**  
 128. **Figure 120**  
 129. **Figure 121**  
 130. **Figure 122**  
 131. **Figure 123**  
 132. **Figure 124**  
 133. **Figure 125**  
 134. **Figure 126**  
 135. **Figure 127**  
 136. **Figure 128**  
 137. **Figure 129**  
 138. **Figure 130**  
 139. **Figure 131**  
 140. **Figure 132**  
 141. **Figure 133**  
 142. **Figure 134**  
 143. **Figure 135**  
 144. **Figure 136**  
 145. **Figure 137**  
 146. **Figure 138**  
 147. **Figure 139**  
 148. **Figure 140**  
 149. **Figure 141**  
 150. **Figure 142**  
 151. **Figure 143**  
 152. **Figure 144**  
 153. **Figure 145**  
 154. **Figure 146**  
 155. **Figure 147**  
 156. **Figure 148**  
 157. **Figure 149**  
 158. **Figure 150**  
 159. **Figure 151**  
 160. **Figure 152**  
 161. **Figure 153**  
 162. **Figure 154**  
 163. **Figure 155**  
 164. **Figure 156**  
 165. **Figure 157**  
 166. **Figure 158**  
 167. **Figure 159**  
 168. **Figure 160**  
 169. **Figure 161**  
 170. **Figure 162**  
 171. **Figure 163**  
 172. **Figure 164**  
 173. **Figure 165**  
 174. **Figure 166**  
 175. **Figure 167**  
 176. **Figure 168**  
 177. **Figure 169**  
 178. **Figure 170**  
 179. **Figure 171**  
 180. **Figure 172**  
 181. **Figure 173**  
 182. **Figure 174**  
 183. **Figure 175**  
 184. **Figure 176**  
 185. **Figure 177**  
 186. **Figure 178**  
 187. **Figure 179**  
 188. **Figure 180**  
 189. **Figure 181**  
 190. **Figure 182**  
 191. **Figure 183**  
 192. **Figure 184**  
 193. **Figure 185**  
 194. **Figure 186**  
 195. **Figure 187**  
 196. **Figure 188**  
 197. **Figure 189**  
 198. **Figure 190**  
 199. **Figure 191**  
 200. **Figure 192**  
 201. **Figure 193**  
 202. **Figure 194**  
 203. **Figure 195**  
 204. **Figure 196**  
 205. **Figure 197**  
 206. **Figure 198**  
 207. **Figure 199**  
 208. **Figure 200**  
 209. **Figure 201**  
 210. **Figure 202**  
 211. **Figure 203**  
 212. **Figure 204**  
 213. **Figure 205**  
 214. **Figure 206**  
 215. **Figure 207**  
 216. **Figure 208**  
 217. **Figure 209**

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84



156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
 301 TTAGGAAGGTTGTAGATGGAATGGGGTGGTCTGTGGTCAACATCTTCTATTTAT 360  
 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAenThrLeuProGlnPheSer 195  
 361 ATATCTGAGACGTCACCACTGAGATTAGGGGACTGTTGAATACGCTACCGCAGTTCACT 420  
 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
 421 GGCCTCGAGGGAGTCTTACTTACTGTATGGTTTCGGATGCTGTTGATGCAATCA 480  
 216 ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
 481 CCTAGCTGGAGATGATGATGCTGGTGCCTTTTCATCCCTTCTCTTCTTCTTCTCCTC 540  
 236 ThrIlePheTyrLeuProGluSerProArgTyrPheValSerLysGlyArgMetAlaGlu 255  
 541 ACGGCTCTTCTTGTGCCCCGAGTCCCCAAGGTGGCTCGTGAGCAAGGTCGAATGCTTGA 600  
 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
 601 GCAAGCGGGTCTTCAGAGACTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 660  
 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
 661 TTGGTTGAGGGTCTTGGAAATGGAGGTGAACCAACCATAGAGGAATATATATTTGGTCCC 720  
 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
 721 GCGATGAAGTACTGATGATCATGATAGTGTGGATAGGATCAATTAAGTTATAT 780  
 316 GlyProGluGluGlnSerTyrIleAlaArgProSerLysGlyProIleMetLeuGly 335  
 781 GGTGCAGAAAGAGGGCTGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 828  
 336 SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354  
 829 AGCACTATGAGTGTGTTGTCGCCATGGAAGTACAATGAGCAGGAGCAAGGCTCAATG 888  
 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374  
 889 ATTGATCTCTTGTCACTGTTGGGAGCGTTCACGAGAGAAGATCCCGACACT---GGA 945  
 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394  
 946 AGCATGAGGAGTGCCTTGTCCACATTTTGGAGTATGTTCACTGTTGGAGGGAATCAA 1005  
 395 AlaLysAsnGluGlnThrAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414  
 1006 CCAAGACATGAAGATTGGATGAAGAAGATCTTGTGGAGAAGGATGATTCATCCATCC 1065  
 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434  
 1066 GAC---CATGGAGATGATCTGAAGATGATCTTAATCTCCGTTGATCTTACGTCAAACG 1122  
 435 ThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSerMet 454  
 1123 ACAAGCATGGAG---AAAGACATGCTCCACACTGCTCATGGAACTCTTCTTACCTTCAGA 1179  
 455 ArgArgGlnThrLeuLeuGlyGluGlyValSerSerThrAspIleGlyGly 474  
 1180 CATGGAAGTCAAGTCAGGAGCTCAAGGGGAAGGAGCGGGTAGTATGGGGATGGAGGT 1239  
 475 GlyTyrGlnLeuAlaTyrIleTyrSerGluLysGlyGluAsnGlyArgLysGluGly 494  
 1240 GGATGGCAAGTGGCATGGAATGAGCGAAAGAGAGATGAATCGGACAGAGAAGAA--- 1296  
 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514  
 1297 -----GAAAGTTTCCAGAGATCTCGACGTTGGCTCAATT 1329

515 ValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 534  
 1330 GTTTCATTCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1386  
 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554  
 1387 TTGGTTAGCAACACAGCTCTTTATTTCCAAAGACCTTCTCAAAGAACATACATATGGTCT 1446  
 555 AlaMetValHisProSerGluValAlaLysGlySerArgTyrLysAspLeuPheGlu 574  
 1447 GCTATGTTACATCCATCCGAA---ACAATAAAGGGTCAATTTGGCATGAATCTTCATGAT 1503  
 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594  
 1504 CTTGGAGTCAAGCGTGCATGATGCTGAGAGTGGACTTCAAAATACCTTCAAGCATCTCA 1563  
 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614  
 1564 GGCATCAACCGGAGTCTTTTACTACACACCGCAAAATCTTTGAGCAGCGGGTGTTCGGGATC 1623  
 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634  
 1624 CTACTATCAACATGGGATTTAGTCTTCTCCTCAGCATCTTACTTATTAAGTGCATTGACA 1683  
 635 ThrLeuLeuLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654  
 1684 ACCTTTGTGATGTTACCTGCAATAGCTGTTCGAATGAGGCTCATGATCTTCTTGTGTCGA 1743  
 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleValVal 674  
 1744 AGGACCTTGTCTCACCACGATACCAATCTGTATGATGATCTCTATTGTTTGTAGTAATC 1803  
 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694  
 1804 TCAATCTTGTTCACATGAACAGCATTTGTCACGGGCTTATCAACGTAAGCGTTGG 1863  
 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714  
 1864 CTCTACTTCTGCTCTCTGATGGTTCGGTCTCGCTCCAAACATCCTCTGTTTCAGAG 1923  
 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734  
 1924 ATTTTCCAACTCGAGTCGGGGAATCTGCATCGCATCTCGCATCTCGCATCTCTGATC 1983  
 735 GlyAspIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754  
 1984 TGTGACATAATCGTCACCTTACAGTCTCCCGCTGCTCAAAATCCATTTGGACTAGCTGCT 2043  
 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774  
 2044 GTGTTGGATGATGCAATCGTATGTTGCAATTCATGGGCTTTGTTGTTTCAATAAGTC 2103  
 775 ProGluThrLysGlyMetProLeuValIleThrGluPhePheAlaValGlyAlaLys 794  
 2104 CCGGAACCTAAAGGATGCCCTTGAAGTCAACAGAGTCTTTCTTCTGTTGGAGCTAGA 2163  
 795 GlnAlaAlaAla 798  
 2164 CAAGCTGAAGCT 2175

## RESULT 12

AK065191

LOCUS

DEFINITION

insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK065191 2368 bp mRNA linear PLN 24-JUL-2003  
 Oryza sativa (japonica cultivar-group) cDNA clone J013002E10, full  
 insert sequence.  
 AK065191  
 AK065191.1 GI:32975209  
 FLI CDNA; CAP trapper.  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.



REFERENCE  
AUTHORS

1. The Rice Full-length cDNA Consortium, National Institute of  
Agrobiological Sciences Rice Full-length cDNA Project Team,  
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
Kiehimoto, N., Yazariki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
Ohtsuka, K., Shishiki, T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K.,  
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
Ikeda, R., Ishibiki, J., Kawanata, M., Yoshimura, A., Miura, J.,  
Kusumegi, T., Oka, M., Ryu, R., Ueda, N., Matsubara, K., KIKEN,  
Kawai, J., Carinici, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M. and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice  
Science 301 (5631), 376-379 (2003)  
22752273  
12869764

Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saichou,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1. .2368
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013002E10"
ORIGIN
Alignment Scores:
Pred. No.:            6.98e-163            Length:            2368
Score:                2583.00            Matches:            508
Percent Similarity:    83.74%            Conservative:        69
Best Local Similarity: 73.73%           Mismatches:        106
Query Match:          63.39%           Indels:             7
DB:                    8                   Gaps:              5

**TITLE**  
**JOURNAL**

**Direct Submission**  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of  
Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp,  
Tel:81-29-838-7007, Fax:81-29-838-7007]  
This clone is one of the 28K full-length cDNA clones from japonica  
rice (*Oryza sativa*).



annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

## FEATURES

```
source
1. .142114
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="10"
/map="R2500"
/clone="OSJNBb0064P21"
/complement(3008..3658)
/genes="OSJNBb0064P21.16"
/notes="predicted by genscan+"
/complement(<3008..>3658)
/genes="OSJNBb0064P21.16"
/complement(3008..3658)
/genes="OSJNBb0064P21.16"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG46103.1"
/db_xref="GI:12039315"
/translation="MVAGDSQARLLVLALIRLLDPAAAAAEPFLFRHSYRATVP
SIVGSDVFWAPESNLTRLHEDLRPTDVLVGLGMLHVLHTDAARYGDALA
SYVDAKSLRSLPVPFPHFWLGLPLLVNHLMTDAKKVHNDTILQAYDVEQRG
LIQRDGGPFLLDVGLSKRGCGGCGGCTAGDGHYDGDYDAVLHMLNALVIESQRI"
complement(3574..3596)
/rpt_family="GC_rich"
/complement(4269..4297)
/rpt_family="(GGGA)n"
5613..5633
/rpt_family="(GGAA)n"
complement(5634..6005)
/genes="OSJNBb0064P21.1"
/notes="identical to histone H4 (H4Cl3) GI:168499 (Zea mays), (Plant Physiol Biochem 25, 235 247 (1987)); EST D23040 from this gene"
complement(<5634..>6005)
/genes="OSJNBb0064P21.1"
complement(5634..5945)
/genes="OSJNBb0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AAG46106.1"
/db_xref="GI:12039318"
/translation="MSGRGKGKGLGKGAKRHKVLRDNIQGITKPAIRRLARRGGV
KRISGLIYBETRGVLKIFLENVIRDAVTYEHARRKTVTAMDVVYALKRQGRITYLGF
G"
complement(7473..7546)
/rpt_family="(CGG)n"
7556..7630
/rpt_family="(GGAGAA)n"
complement(8238..8259)
/rpt_family="AT_rich"
complement(8915..8963)
/rpt_family="AT_rich"
complement(11055..11097)
/rpt_family="AT_rich"
complement(11244..11264)
/rpt_family="AT_rich"
14313..18336
/genes="OSJNBb0064P21.2"
/notes="identical to calcium dependent protein kinase GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993)); EST C22538, C22540, AA750452, C99271, C22539 from this gene"
join(14313..15004,15338..15481,15564..15716,15894..16009,16646..16810,17112..17333,17903..18336)
/genes="OSJNBb0064P21.2"
```

## CDS

```
join(14323..15004,15338..15481,15564..15716,15894..16009,16646..16810,17112..17333,17903..18025)
/genes="OSJNBb0064P21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AAG46110.1"
/db_xref="GI:12039322"
/translation="MNSCQGTGYNNYQNSRPNQDRFASRYVDGNDTDCYSGSSR
ASLAGALRQGLNLSKSPVGLGTPNRELYTLRGELGGQFGKTYLCTEISTGQYACK
TILKNSLRCSIDEDVRREIQIMHLSGQKNITVTKDYEDQAVHIVMELCAGSELF
SKIQRGHYSERKAAELIKIIVGIIETCHSHGVMRDLKPENFLLDADDEFSVKAI
D FGLSVFERPGQVREVVGSPYIAPEVLEKRYGPEADIWTAGVLLVLLTGVPPFWAD
TQSGIYEKVLGDRIDPKSNRPRI SDSAKDLIKMLCPYPSERLKAHEVLKHPWICDN
GVATNRALDPSVLPRLKQFSANRRLKKLSLQIIAERLSEEEIVGLREMFKAMDKNRS
VVTGELKGLKRYSSVFKDTEINLMEAADDTITINBEEFTAAAVSLNKIERKHLM
AAFTYFDKDGSGFITVDKLQAKMERNMEDTFLFLEMLEVQDNDQOIDYAEFVTWQ
SNNFGLGQTVESLNVALREAPQVY"
complement(18675..18732)
/rpt_family="GC_rich"
/complement(20605..20634)
/rpt_family="AT_rich"
22393..24891
/genes="OSJNBb0064P21.6"
/notes="predicted by genemark hmm"
<22393..>24891
/genes="OSJNBb0064P21.6"
22393..24891
/genes="OSJNBb0064P21.6"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG46113.1"
/db_xref="GI:12039325"
/translation="MAGSGEAMASGLDLGBYLDRPDIAHRRRAASVAIVRGGGDP
RIVDGRDRRARSRRLSSWRRPAGGETSVESRRYGTAAAAAEEVGKTRAAAIAW
BKKPVRLSRIGKRCGLLSVEVAGVGPASMDGLRVAIVRKAEIKDGAQVTMPA
TVRGCADEFVETLVRCNILYFAGAGTGKPLKLEPRFVVVPAEPGVRLCAHAVD
VSSLLDSLSQKSGSEGRVRWPDATVTLGKATGGELLKLGQLMDAGLCLYQAA
EKVDVVVDSPALARAHNKNSFVARTSGPKLSASDAALSPSMRAIKQLIDRLSVDEH
GDPVTSLI PRKLADDELSDGVLPEYVVVDKGTVEVYVHHAHRDLKELDSIAEQ
IBATILAMTNGGKSGPPSPKVDQKQCLDADBEEMTVFEFLRKLVEVDDKGRKLQPM
PSESEKAAAAAPPVVPDLGPGLTAVQTRDGGFLVSMNPFDLPLEKGDAPPKLQAMQ
VSRPVLPSMAATGFDVLQKMAAGGAEVDRKVARLGMDNLTKTPEQVFEGLTA
EAVIGRRTEGASSNAARSVRLVRLKLAASVDSGRSEFVATGWTAADDPETLEVLTA
SIQKLEMAVDALMTQAENADEDAPEFVAPAAAGDATTVFDALVSPSEWSESQSGRV
TVVAAIQLRPSRSEYEAAGPVAVVQSARLIGAAGNSGGRFKVRSLSHVGVOQLRCPV
GGAGGGRASWGAERQKLTAMQWALAHGPARAAGRPARPTTTTTPPSQARQRPDVVM
SLSSRLAGMWLKTVRNPDVRVCGATAAGSGGCGN"
complement(23766..23788)
/rpt_family="GC_rich"
complement(24616..24639)
/rpt_family="GC_rich"
complement(24693..24714)
/rpt_family="GC_rich"
24853..24903
/rpt_family="(CGG)n"
complement(26149..26194)
/rpt_family="AT_rich"
complement(26861..26905)
/rpt_family="AT_rich"
complement(27502..27537)
/rpt_family="GC_rich"
complement(27628..27663)
/rpt_family="AT_rich"
complement(28100..28132)
/rpt_family="AT_rich"
complement(29892..29929)
/rpt_family="(TA)n"
29893..29928
/notes="with dinucleotide (AT) repeat (35 to 42 copies)"
34449..38345
/genes="OSJNBb0064P21.3"
/notes="similar to sugar transporter GI:1495273 (Arabidopsis thaliana); EST AA231788, D48171, C72446,
```



[illegible]

Db	37861	CATTACAGTTTCCTT-CTCCTTGCACAATCCCTATCTCGTAGTAGTACTAGCTATCTTTG	379131
Qy	673	ValValSerAsnLeuileAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer	692
Db	37920	ATTCTGGTCAATATTCTGGATGTGGGACCATGGTTTCATGCCTCACTGTCCACAGTCACT	37979
Qy	693	ValleValTyrrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys	712
Db	37980	GTCACTACTCTACTTCTGCTCTTCTTGTGCATGGGGTTCGGGCTTATTCCAAACATCTCTGT	38039
Qy	713	AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleAlaPheThrPhe	732
Db	38040	GCAGAGATTTTCCGACACACCGTTGTGGCATCTGCATAGCCATCTGCCCCTACATTC	38099
Qy	733	TrpIleGlyAspIleIleValThrTyrrSerLeuProValMetLeuAsnAlaIleGlyLeu	752
Db	38100	TGGATCGGTGATATCATTTGTGACATACACCCCTCCCGTGATGCTCAACGCCATTGGACTC	38159
Qy	753	AlaGlyValPheSerIleTyrrAlaValValCysLeuIleSerPheValPheValPheLeu	772
Db	38160	GCTGGAGTGTGGATCTPACGCAGTGGTCTGCATCTGGCTTCTCTCTCTCTCATG	38219
Qy	773	LyavalProGluThrTyrrGlyMetProLeuGluValIleThrGluPhePheAlaValGly	792
Db	38220	AAGTCCCGAGCAAGGCGCATGCTCTTGAAGTCATCACCAGATTCTTCTCTGTGCGA	38279
Qy	793	AlaIleGlnAla 796	
Db	38280	GCAAGCAGGCC 38291	
RESULT 14			
AE017116			
LOCUS	AE017116	300957 bp	linear
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 10, section 70 of		linear PLN 06-JUN-2003
ACCESSION	AE017116		
VERSION	AE017116		
KEYWORDS	AE017116.1	GI:31433303	
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Oryza sativa (japonica cultivar-group)		
AUTHORS	Oryza sativa (japonica cultivar-group)		
CONSTRM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
JOURNAL	Ehrhartoideae; Oryzaceae; Oryza.		
COMMENT	1 (bases 1 to 300957)		
	The Rice Chromosome 10 Sequencing Consortium		
	In-depth view of structure, activity, and evolution of rice		
	chromosome 10		
	Science 300, 1566-1569 (2003)		
	2 (bases 1 to 300957)		
	Buell, C.R., Wang, R.A., McCombie, W.R., Messing, J. and Yuan, Q.		
	Direct Submission		
	Submitted (05-MAY-2003) The Institute for Genomic Research, 9712		
	Medical Center Dr, Rockville, MD 20850, USA		
	This is the pseudomolecule for rice chromosome 10, which was		
	constructed by resolving discrepancies between overlapping BACs,		
	trimming the overlap regions, and linking the unique sequences to		
	form a contiguous sequence. Genes in individual BAC clone were		
	identified by a combination of several methods: Gene prediction		
	programs, searches of the complete sequence against a peptide		
	database and EST databases. Genes with similarity to other proteins		
	are named after the database hits. Genes without significant		
	peptide similarity but with EST similarity are named as unknown		
	proteins. Genes without protein or EST similarity, that are		
	predicted by more than two gene prediction programs over most of		
	their length are annotated as hypothetical proteins. Genes		
	encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy,		
	http://genome.wustl.edu/eddy/tRNAscan-SE/).		
FEATURES	Location/Qualifiers		
source	1. 300957		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		
	/cultivar="Nipponbare"		

## FEATURES

```

1. .300957
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"

```

```

/db_xref="taxon:39947"
/chromosomes="10"
<1..12488
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBa0040D23
(GB:AC074196)."
<3310..>4569
/locus_tag="OSJNBa0040D23.23"
/note="similar to chloroplast
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<3310..>4569
/locus_tag="OSJNBa0040D23.23"
3310..4569
/locus_tag="OSJNBa0040D23.23"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54833.1"
/db_xref="GI:31433304"
/translation="MGRILAPMQLLVLCISVTTCAAHGLRGLDQQQMRGRIILADAT
AAPGGAVVFLHWSGAHYVNFITGTPPOAVSGIDVLSGELVWTQCAACSSGCFKQE
LPVDFPSASVTRAAQCCSPICKSIPTKNCSDGECGEYAPSMFGDTFGIATSDAIAI
GNAEGRILAFCCVYADSGIDGAMDGPFGVGLRTPMSLVQSNVTAFSYCLAHGPG
KXSLFLGASAKLAGAKSNPTPLLGQHASNTSDGSDPYTYVLEGGIKAGDVAVA
ASSGGGATVLETFRLSYLPDAAVQALKEKVTVAALGSPMANPEPPFDLCQVAAA
VSGVFDLVFTFOGGATLTAQPSKYLGGNGNGTVCILSILSTRLDSADDGVSTLSL
LQENVHFLFDLKEKTLSEFAPDCSLT"
<5409..>5992
/locus_tag="OSJNBa0040D23.24"
/note="predicted by Ignessh"
join(<5409..5547,5811..>5992)
/locus_tag="OSJNBa0040D23.24"
join(5409..5547,5811..5992)
/locus_tag="OSJNBa0040D23.24"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54836.1"
/db_xref="GI:31433305"
/translation="MTGNLSNRICLTKTPCILKHPTCGANMLRORTSTKAASHCP
TGKRSKSDSLISAVLONQNLHGQTKAPLSTIIIVARGSTFVARRNFPFKTKRKN
MYAL"
<8062..>9183
/locus_tag="OSJNBa0040D23.25"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<8062..>9183
/locus_tag="OSJNBa0040D23.25"
8062..9183
/locus_tag="OSJNBa0040D23.25"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54835.1"
/db_xref="GI:31433306"
/translation="MRGRLLADAAGGGGVVPIYWSQPLYMANLTIGTPPOPASAI
THLAGEFVWTCSPCCFKQDLPLFNPESASTYRPEPCQTCALCESVPASTCGSDGVC
SYEVTFMGDSGIGGTTTFAIGTATASLAFGCAMDNIKOLLGASGVGLRTPMSL
VQOMNATPSYCLAPHGAKGKSALLIGASAKLAGKSAATPVNTSDSSDMHL
RGIKGQDVIILAPPNGSVLVDTLFCVSFVDANAFQAKKAVTVAVGNAAPATKFP
DLCFPAAAGANSLSLFPDVLITFGQAALTVFPFSKMTDAGNGTIVCLAMWSSAML
NUTTBLSILGRHQHNIHFLDLDKETSLEFAPDCSSL"
<11356..>12573
/locus_tag="OSJNBa0025B05.4"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<11356..>12573
/locus_tag="OSJNBa0025B05.4"
11356..12573
/locus_tag="OSJNBa0025B05.4"
/codon_start=1
/product="putative chloroplast nucleoid DNA binding
protein"
/protein_id="AAP54836.1"
/db_xref="GI:31433307"
/translation="MGRPVATLFLVLCFTSVCTSLSEQATRGRLLAGVDATPPAAGCA

```

```

VAVPIYLSQGLYVANFTIGTPQPVSAVVDLTGBLVWVWTQCTPCQPCFEQDLPLPDPIT
KSTSTFRGLPCGSHLCESIPESNRCTSDVCIYEAFPTKAGDTGGKAGTDTFAIGAAKET
LFGCVWMTDKRLKTIIGSGIVGLRTPMSLVQSNVTAFSYCLAAKSSGALFLGAT
AKQLAGKNSPTFFVITKSAGSDNSGPNPKYKLAGIKTGGAPlQAASGSGSTVLLD
TVSRASYLADGAYKALKAL/TAAGVQVPVAGPYPKDYDLCTFPKAVAGDAPELVFTPDGG
AALTVPANYYLLASNGTVCILTIGSSASINLTGLEGASILGSLQQENVHVLFDLKEE
TLSFKPADCSLT"
12489..23141
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBa0025B05
(GB:AC096782)."
complement(<16252..317307)
/locus_tag="OSJNBa0025B05.3"
/note="predicted by Ignessh"
complement(<16252..317307)
/locus_tag="OSJNBa0025B05.3"
complement(16252..17307)
/locus_tag="OSJNBa0025B05.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54837.1"
/db_xref="GI:31433308"
/translation="MVSAAAQGVVACVVLFPVPMGLAGHLSRNKVLFFSGALFVSL
AVGVHLSPLSLPLHLLAASFPHPHPTSSSSSSAASSSCVPLLRHVSADAGGEGV
GRAMWPSLASTCGLRLARDLSDLLNGSWVMVAGSQARLLVALLRLLLDPAA
AAAEPLFRHSIDYRATVPARGISVDFVWAPFESNLTLLHEDLRLPRTPDVFLVGS
GLWMHVTDAARYGDALASVVDRAKSLRSLPVPFPMFWLGLPLVNMHMLTDAKK
VHMNTTILQAVDLEVEQRGLLQRDGGPFLLLDVGKLSRGCGQCTADGMHYDGVYDA
VLHIMNALVIESQORI"
complement(<18878..19249)
/locus_tag="OSJNBa0064P21.1"
/note="similar to histone H4 (H4C13) GI:168499 (Zea mays),
(Plant Physiol Biochem 25, 235 247 (1987)); EST D23040
from this gene"
complement(<18878..19249)
/locus_tag="OSJNBa0064P21.1"
complement(18878..19189)
/locus_tag="OSJNBa0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AAP54838.1"
/db_xref="GI:31433309"
/translation="MSGRGKGKGLGKGAKHRKRVLRDNTOGITKPAIRLRARGGV
KISGLIYEETRGVLKIFLENVIRDAVTYTHARRKVTAMDVVYALKRQGTLYLFG
G"
<19395..>20598
/locus_tag="OSJNBa0025B05.2"
/note="similar to B regulatory subunit of PP2A
GB:AAB58900 GI:2160690 (Arabidopsis thaliana)"
join(<19395..19473,19781..19910,19992..>20598)
/locus_tag="OSJNBa0025B05.2"
join(19395..19473,19781..19910,19992..20598)
/locus_tag="OSJNBa0025B05.2"
/codon_start=1
/product="putative protein phosphatase 2A subunit"
/protein_id="AAP54839.1"
/db_xref="GI:31433310"
/translation="MTWLVTTGBRDAIRGRVVLHRWMDGAGPPGGAGGAGDGRROL
VPASQCRPFRPRRRPTCSSTRILRNHVDSEFLACLVAFEDSPERDLKYV
YHQLYSKUTVERAFMORIMAAALNFVDYDTSFEAEHRHCGVGELEICGSLINGFAYP
LKEHRMARVLLPLHRTVWHTYHRLQSLTCSGIRSRHLLPFPPLAGQMPFGGAHR
RARGRRDPLSAPGLPAAASAAAHRFVCFHFNHRRHFSOKKIDFIYICTFN"
23142..155358
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBb0064P21
(GB:AC073166)."
<27557..>31580
/locus_tag="OSJNBb0064P21.2"
/note="similar to calcium dependent protein kinase
GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993));
EST C22538, C22540, AA750452, C99271, C22539 from this
gene"
join(27557..28248,28582..28725,28808..28960,29138..29253,

```

29890. 30054,30356..30577,31147..31580)  
 /locus\_tag="OSJNB0064P21.2"  
 join(27567..28248,28582..28725,28808..28960,29138..29253,  
 29890..30054,30356..30577,31147..31269)  
 /locus\_tag="OSJNB0064P21.2"  
 /codon\_start=1  
 /products="calcium-dependent protein kinase"  
 /protein\_id="AAP54840.1"  
 /db\_xref="GI:31433311"  
 /translations="MGNSCQGTGYNNYQNSRNFQDRFASRYVDGNDTDCYSGSSR  
 ASLAGLRQGLNLKSPVLGYKTPNVRRELYTLGRLEGQGFQKTYLCTEISTGCGQYACK  
 TILKSLNRQVSDIDVREIQIMVHLISGQKNIVTIKDYEDQAVHIVMELCAGGELF  
 SKIQKRHSYERKAAELIKIIVGIIETCHSHGVMHRDLKPENFLLDADDFSVKALD

## Alignment Scores:

Pred. No.: 3.13e-155 Length: 300957  
 Score: 2514.00 Matches: 552  
 Percent Similarity: 53.77% Conservative: 75  
 Best Local Similarity: 47.34% Mismatches: 109  
 Query Match: 61.69% Indels: 431  
 DB: 8 Gaps: 9

US-10-051-909-32 (1-800) x AE017116 (1-300957)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75  
 Db 48060 ATGCGGGCCGCGTGGTGGCCATCGCGCCCTCCATCGGCACTTTGCTGCAGGGCTGG 48119  
 QY 76 AspAsnAlaThrIleAla----- 81  
 Db 48120 GATAATCAACCATTTGCAGGTAAAGTTCGTGGTTAGAGAAAAACGAATCTTTTTTTTCT 48179  
 QY 81 ----- 81  
 Db 48180 GCCATAGTTTGTGGCCATCGTTATGTGACGCTAGGGCCCGTGTGGTGGCAATTAT 48239  
 QY 81 ----- 81  
 Db 48240 TGTGATACTTGTGGAGATCATGATGGATTTGGTGGAAATCATCTGCAGTAAAAATAGTTTT 48299  
 QY 81 ----- 81  
 Db 48300 AGTTGTTCTTCTGTTTCATGATATCAAAAAAACAACGTGATATTTTTTTTCAGTA 48359  
 QY 81 ----- 81  
 Db 48360 GAATTTTGATCAATTATGATAGTAGATTAAACAAATGTACGGATCTCATGAATCATCTGT 48419  
 QY 81 ----- 81  
 Db 48420 ATTTATGGAACCTTGATACTGCTGCATCGTAAAGGTTTCAGAAAAATTTGATTGATACCGATG 48479  
 QY 82 -----AlaIleValle 85  
 Db 48480 GTTTTGACTTCTGCTGACCCCTGCGACTTGTGGGATGATCTTTATCAGGTGCGGTACT 48539  
 QY 85 uTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValse 105  
 Db 48540 GTACATCAAGAGGAATTAACATTGCAGAGGAGCCCTTATCGAAGCCCTGATCGTGGC 48599  
 QY 105 rMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerIle 125  
 Db 48600 CATGTCGCTCATTTGGGCGCAGCATCATCAGACGTTCTCTGGAGCAGTGGCTGATCTTT 48659  
 QY 125 eGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMe 145  
 Db 48660 TGGTAGCGCGCCCATGCTGATCGCGTGGCTCTCTACTTTGTTAGTGGGCTAGTGAT 48719  
 QY 145 tLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIle 165  
 Db 48720 GCTTTGGCGCCAAATGTGATGTGTGCTCTTGGCGAGGCTCATACCGGGTTCCGGAT 48779  
 QY 165 eGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAr 185

Db 48780 CGGTTTGGCTGTCCAGCTTGATACATTGTACATCTCTGAGACTGCCCCGAGCATCTAG 48839  
 QY 185 gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCy 205  
 Db 48840 AGGACTGCTAAACACACGCTGCCGAGTTCAGTGGGTCTGGAGGATGTTCCTTTTCATCTG 48899  
 QY 205 sMetValPheGlyMetSerLeuSerProSerProAspTyrArgIleMetLeuGlyValle 225  
 Db 48900 CATGGTATTTGGGATGTCCCTCATGCCACAGCCAGATTGAGGATCATGTGTGCGTCT 48959  
 QY 225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProAr 245  
 Db 48960 ATCAATACCATCTTATATACATTGCAATTGACATCTTTTACTTACCTGATGCCGAG 49019  
 QY 245 gTyrLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgG 265  
 Db 49020 GTGGCTCGTGAGCAAGGAAGAAATGCTGTAGGCCCAAGCGTGTGTGCAAGCCCTGCCGTG 49079  
 QY 265 YLysAspAspValSer----- 270  
 Db 49080 AAGAGAAGATGTTTTCAGGTTTGTATGTCTTACCAGCATTTTTTCCCTGTGAGGATCTAGA 49139  
 QY 270 ----- 270  
 Db 49140 AAGTGAGAAACTATTTAATGTCTATCAGTATAGGTATATTATCTATGCGATGGAGTTTCT 49199  
 QY 271 -----GlyGluLeuSerLeuLeuGluGlyLeuGluValGlyG 284  
 Db 49200 CATGTTGAATTAACCTGTCAGGAGAAATGCGCCCTTCTCGTTGAAGGTCTCGGGGTGGGA 49259  
 QY 284 lyAspThrSerIleGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuV 304  
 Db 49260 AAGACACAAAAATTGAGGAATACATAATTGACCTGATGATAGCTTGTCTGATGAAGGC 49319  
 QY 304 alThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpI 324  
 Db 49320 TGCTCTCA--GATCCAGAGAAAGTCAAACTGTATGCTCTGAGGAAGGCTTATCGTGGG 49376  
 QY 324 leAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH 344  
 Db 49377 TTGCCCGTCTGTTCCAGGCAAAAGTGCATCTGGAAGTGCATTAGGTCTCATCTCTGTC 49436  
 QY 344 isGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlyS 364  
 Db 49437 ATGCTAGTATGTCAGTACGGGTAAGCCCTTGTGGATCCTGTGTCCACCTTTTGGAA 49496  
 QY 364 erValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnP 384  
 Db 49497 GTGTCCATGAGAAAGATGCCTGAGATAATGGGAAGCATGCGGAGCACATTTCTTCTAACT 49556  
 QY 384 heGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluA 404  
 Db 49557 TTGGCAGCATGTTTAGTGTGGCGGAACAGCAGCAAGCTTAAGGTGATTGGGATCTCTGAGA 49616  
 QY 404 snLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspA 424  
 Db 49617 GT---CAACGGGAGGTGAAGATTATGGATCAGACCATGCTGGGATGACATTTGAAGATA 49673  
 QY 424 snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---V 443  
 Db 49674 GCCTCAAAAGCCCACTATTCTCTGTCAGCGCAAGCGTGGGAAGAAAGAGATCGCTG 49733  
 QY 443 alHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluG 463  
 Db 49734 CACTCATGGCAGTATAAATGGGTGCTGTGGGA-----AGAAAGTAGTAGTCTCATCGAG 49787  
 QY 463 lyGlyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrps 483  
 Db 49788 GCGGGGAGGAGTAAAGCAGCATCGGCATTTGTTGGGGATGGCAGTTGGCTTTGGAATGGA 49847  
 QY 483 erGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisG 503





```
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
173..1890
/gene="F23E12.10"
CDS
join(173..374,456..919,1089..1610,1678..1890)
/gene="F23E12.10"
/codon_start=1
/product="hypothetical protein (fragment)"
/protein_id="CAAL8726.1"
/db_xref="GI:3080407"
/db_xref="GOA:O65486"
/db_xref="UniProt/TREMBL:O65486"
/translations="FNAAVAGCFFAVESVLWPSSTDSSTSLPNTTSMVILSAVTASV
VSEIGLSEPAFVPDYDRSPGELPLYLGLGALCVSLALSRCTSSMTSAVDSLNVK
DAGIPKAVPVMGSLVGIIALYVPEVLYWGFQNVILLLEKRPVKGSLADLLQLVA
VKIAAWACASGLVGYAPSLFIGAAGMAYGKFIGLALAQNPDPFNLSILEVASFPQ
AYGLVGAATLACGVCQPLTAVLLFELTQDYRIVLPGLGAVGMSWITGGQSKRQET
RETKRKRSQAVSLTSSDDSESTNNLCEVESLCLDDSLNQSEELPKSIFVSEA
MRTRFATVMSTSLSEALTRMLTEKOSCALIVDPDNIFLGILTLSDIQFSPKARKEGN
NRPKDIFVNDICSRSGCKKVPVTVPDMLLAQTIMNKHLSHVAVVGSDAPRI
HPVGVDREICITLRR"
173..374
/gene="F23E12.10"
/intron
375..455
/gene="F23E12.10"
/exon
456..919
/gene="F23E12.10"
/intron
920..1088
/gene="F23E12.10"
/exon
1089..1610
/gene="F23E12.10"
/intron
1611..1677
/gene="F23E12.10"
/exon
1678..1890
/gene="F23E12.10"
/gene="F23E12.30"
3124..4428
/gene="F23E12.20"
CDS
join(3124..3241,3323..3465,3595..3696,3945..4107,
4211..4428)
/notes="similarity to dihydroflavonol-4-reductase, Gerbera
x sp., PIR2:S35189"
/codon_start=1
/product="putative protein"
/protein_id="CAAL8727.1"
/db_xref="GI:3080408"
/db_xref="UniProt/TREMBL:O65487"
/translations="MDQAKGKVCVTGASGFLASMLVKRLLEGVEVIGTVDRDPGNEKK
LAHLWLEGAKEKRLVKAIDLMEGSPDNAINGCQGVPHFTASPEILRPALEGTLNVLK
SCRKNPSLRKRVLTSSSTVYALSKTALAEAAWKFSENGIDLVTLVPLSPFLVGPSPID
PDLCSSTASDVLGKGTETKEFKQHWGQGYHIDDVARTHVVPFEHAAQGRYICSSNV
ISLEELVSLSARYPSLPIPKRFDTPLN"
3124..3241
/gene="F23E12.20"
/intron
3242..3322
/gene="F23E12.20"
/exon
3323..3465
/gene="F23E12.20"
/intron
3466..3594
/gene="F23E12.20"
```

```
/number=2
3595..3696
/gene="F23E12.20"
/number=3
3697..3944
/gene="F23E12.20"
/number=3
3945..4107
/gene="F23E12.20"
/number=4
4108..4210
/gene="F23E12.20"
/number=4
4211..4428
/gene="F23E12.20"
/number=5
4939..6163
/gene="F23E12.30"
complement(join(4939..5013,5109..5195,5307..5342,
5451..5651,5725..5811,6161..6163))
/gene="F23E12.30"
complement(join(4939..5013,5109..5195,5307..5342,
5451..5651,5725..5811,6161..6163))
/notes="Contains Clathrin adaptor complexes small chain
signature [VYKRYASLYF]
contains EST gb:T44219"
/codon_start=1
/product="clathrin assembly protein AP19 homolog"
/protein_id="CAAL8728.1"
/db_xref="GI:3080409"
/db_xref="GOA:O23685"
/db_xref="UniProt/TREMBL:O23685"
/translations="MIHFVLLVSRQKVLTKWYSPYAKERSKVIRESGVILNRGP
KLICNFVEWRYGVVYKRYASLYFCMCIDQSDNELEVLIIHHVVEILDYRFGSVCELD
LIFNFHKAYIILDELLIAGELQESSKKTARIISAQDLVEVAKKEASSISNIIAQT
NR"
complement(4939..5013)
/gene="F23E12.30"
/number=1
complement(5014..5108)
/gene="F23E12.30"
/number=1
complement(5109..5195)
/gene="F23E12.30"
/number=2
complement(5196..5306)
/gene="F23E12.30"
/number=2
complement(5307..5342)
/gene="F23E12.30"
/number=3
complement(5343..5450)
/gene="F23E12.30"
/number=3
complement(5451..5651)
/gene="F23E12.30"
/number=4
complement(5652..5724)
/gene="F23E12.30"
/number=4
complement(5725..5811)
/gene="F23E12.30"
/number=5
complement(5812..6160)
/gene="F23E12.30"
/number=5
complement(6161..6163)
/gene="F23E12.30"
/number=6
7352..7582
/gene="F23E12.40"
complement(7352..7582)
```

CDS	/gene="F23E12.40"		Db	40189	ACCTGAGATTAGGGGACTGTTGAATACGCTACCGCAGTTTCACTGGCTCTCGAGGGATGTT	40248	
	complement(7352..7582)			Qy	201	eLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTyrArgIleMe	221
	/gene="F23E12.40"			Db	40249	CTTATCTTACTGATGTTTTCGGAATGCTGTTGATGCCATCACCTAGCTGAGATTGAT	40308
	/codon_start=1			Qy	221	tLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuPr	241
exon	/protein_id="CAA18729.1"		Db	40309	GCTTGGTGTCCTTTTCATCCCTTCCCTTCTTTTCTTCTCACGGTCTTCTTCTG	40368	
	/db_xref="GI:3080410"		Qy	241	oGluSerProArgTyrLeuValSerIlyGlyArgMetAlaGluAlaLysValLeuGl	261	
	/translation="MPLVRKNGVMRIGSPMLHSENLVGDGVEAPPGRGERDEIGEEV"		Db	40369	CGAGTCCCCAAGGTGCTGCTGAGCAAGTCAATGCTTGAAGCAAGCGGGTCTTCA	40428	
	complement(7352..7582)		Qy	261	nLysLeuArgGlyLysAspValSer	270	
gene	8233..9045		Db	40429	GAGACTGCGTGGTGGCGAAGATGTCTGTTAGTCTTAAATCGTTCGACTCTAATTC	40488	
gene	/gene="F23E12.50"		Qy	271	-----	274	
CDS	/gene="F23E12.50"		Db	40489	TTCAATGAACCAATCTTTTCTATTGAATGTTCTTATGTGGCAATGTATAGGTGAGATGG	40548	
	complement(8233..9045)		Qy	274	erLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleleG	294	
	/gene="F23E12.50"		Db	40549	CTTTGTTGGTGGAGGCTTTGGAATTTGGAGGTGAAACCAACCATAGAGGAATATAAATG	40608	
	/notes="similarity to DNA-binding protein PDI, Pisum sativum"		Qy	294	lyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGluIleThrL	314	
Alignment Scores:	/codon_start=1		Db	40609	GTCCCGGGATGAAGTTACTGATGATCATGATAGCTGTGGATAAGGATCAATTAAGT	40668	
	/product="putative protein"		Qy	314	eutyGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetL	334	
	/protein_id="CAA18730.1"		Db	40669	TATATGGTGCAGAAAGGCTGAGTTGGTGTCTAGGCCAGTCAAAAGG	40718	
	Length: 86710		Qy	334	eGlySerValLeuSerLeuAlaSerArgHisGlySer	353	
Pred. No.:	3.27e-155	Matches: 523	Db	40719	-----GGAACCACTATGAGTGTGTTGTCCTCCATCGAAGTACAATGAGCAGGAGCAAGCT	40776	
Score:	2502.00	Mismatches: 91	Qy	353	roLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaG	373	
Percent Similarity:	68.91%	Indels: 153	Db	40777	CATTGATTGATCCTCTTGTCCACTGTTTGGAGCGCTTCACGAGAAGATGCCGACACT	40835	
Best Local Similarity:	58.70%	Gaps: 12	Qy	373	lyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspG	393	
Query Match:	61.40%		Db	40836	-----GGAAGCATGAGGAGTGCCCTGTTGCCACATTTTGGGAGTATGTTCACTGTTGGAGGA	40893	
DB:	8		Qy	393	lnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrA	413	
US-10-051-909-32 (1-800) x ATP23E12 (1-86710)		Db	40894	ATCAACCAAGACATGAAGATTGGGATGAAGAGAAATCTTGTGGAGAAGGTGAGGATTATC	40953		
Qy	52 LeuGluAspLys---MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70		Qy	413	laSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgG	433	
Db	39649 ATAGAAGACCGAATTATGATGGAGCTGCTGTTGTTGCTATGCTGCTGTGGCAAC 39708		Db	40954	CATCCGAC---CATGGAGATGATTCCTGAAGATGATCTTCAATCTCCGTTGATCTCACGTC	41010	
Qy	71 LeuLeuGlnGlyTrpAspAsnAlaThrIle- 80		Qy	433	lnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuS	453	
Db	39709 TTGTTACAAGGATGGGATACGCAACTATTTCGAGGTGATTGATTAATCTTCTGTATCCATA 39768		Db	41011	AAACGACACAGCATCGAG---AAAGACATGCTCACACTGCTCATGTGGAACCTCTTTACCT	41067	
Qy	80 ----- 80		Qy	453	erMetArgArgGlnThrLeuLeuGlyGlyAspGlyValSerSerThrAspIleG	473	
Db	39769 ATTCTGTTTATGAATTTAGGATCGATAAAGTCAGGATTTTATGGATACAAATTCACAGA 39828		Db	41068	TCAGACATGGAAGTCAAGTCCAGGAGCTCAAGGGAAGGAGCGGTAGTAGGGGATTG	41127	
Qy	81 -----Al 81		Qy	473	lyGlyGlyTrpGlnLeuAlaTrpIlySerGlyLysGluGlyGluAsnGlyArgLysG	493	
Db	39829 TCITTTAGCTATCTCTTTGCTGTGTGATGGATTTTAAGTTTCTTTGGGTACTCTGC 39888		Db	41128	GAGGTGGATGCGCATGTCATGGAATGGACGGAAGAGAGATGAATCGGACAGAAAG	41187	
Qy	81 aAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGl 101		Qy	493	luGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlyS	513	
Db	39889 AGGAGCTGTGTGTACATAAAAGAGGATTTAATTGGAGAGTAATCCATCAGTGGNAGG 39948		Db	41188	AAGTGGGTTTTAAACGGATATCTTGCATCAAGAGGTTTCCAGGATCTCCAGCTGGCT	41247	
Qy	101 yLeuileValSerMetSerLeuileGlyAlaThrIleValThrPheSerGlyProLe 121		Qy	513	erIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaA	533	
Db	39949 TCTAATTGTGGCGATGTCACCTATTGTGTGCTACTCTGATTACAAACATGCTCTGGAGGGT 40008		Qy	41248	CAATTGTTTCTTGCCTGGTGTGATGGAAACCGGTGAG---GCAGATTTTGTACAGCGT	41304	
Qy	121 userAspSerIleGlyArgArgProMetLeuileLeuSerSerIleLeuTyrPhePheSe 141		Db				
Db	40009 AGCTGATTGGCTGTGCTGCCGCTCCCATGCTAATATTGCTCCTCAATCTCTACTTTGTG 40068		Qy				
Qy	141 rGlyLeuileMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAs 161		Db				
Db	40069 TTCTCTAGTATGCTATGCTCCGAATGTTTATGTTGTTGCTCTTAGGAAGGTTGTTAGA 40128		Qy				
Qy	161 pGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluileAlaPr 181		Db				
Db	40129 TGGATTTGGGGTGGTCTTGTGGTCACACTGTGTTCTTATATATATATCTCAGAGTCGACC 40188		Qy				
Qy	181 oSerGluileArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPh 201		Db				



**This page blank (uspio,**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 12:06:34 ; Search time 923.577 Seconds

(without alignments)  
5127.663 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSLAVQPTPTDPLDRR.....PLEVITEFPAVGAQAQAAKA 800

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US10051909/runat\_13042005\_074037\_14040/app\_query.fasta\_1.1678  
-DB=N\_Geneseq\_16Dec04 -OPMT=fastap -SUFFIX=rng -MINWATCH=0\_1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USPR=US10051909 @CGN 1 1 891 @runat\_13042005\_074037\_14040 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIO  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1980s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4075	100.0	2777	12 ADG47935	Adg47935 Corn Arab
2	4059	99.6	2908	12 ADM47932	Adm47932 Polynucle
3	3517	86.3	2824	6 ABK51962	Abk51962 Corn cont
4	3517	86.3	2824	8 ABX93198	Abx93198 cDNA enco
5	3517	86.3	2824	12 ADG47905	Adg47905 Corn Arab

6	2674	65.6	2601	6	ABK51965	Abk51965 Soybean c
7	2674	65.6	2601	8	ABX93201	Abx93201 cDNA enco
8	2674	65.6	2601	12	ADG47911	Adg47911 Corn Arab
9	2584	63.4	2190	6	ABZ13510	Abz13510 Arabidopo
10	2275	55.8	2205	6	ABZ14449	Abz14449 Arabidopo
11	2275	55.8	2205	6	ADG87978	Adg87978 A. thalia
12	1670.5	41.0	1692	6	ABK51966	Abk51966 Soybean c
13	1670.5	41.0	1692	8	ABX93202	Abx93202 cDNA enco
14	1670.5	41.0	1692	12	ADG47913	Adg47913 Soybean A
15	1424	34.9	1487	6	ABK51968	Abk51968 Wheat con
16	1424	34.9	1487	8	ABX93204	Abx93204 cDNA enco
17	1424	34.9	1487	12	ADG47917	Adg47917 Wheat Ara
18	1035	25.4	1412	12	ADJ39747	Adj39747 Plant cDN
19	943	23.1	1009	6	ABK51969	Abk51969 Wheat cDN
20	943	23.1	1009	8	ABX93205	Abx93205 cDNA enco
21	943	23.1	1009	12	ADG47919	Adg47919 Wheat Ara
22	620	15.2	751	12	ADJ41683	Adj41683 Plant cDN
23	616	15.1	1806	10	ADC07791	Adc07791 Rice DNA
24	600	14.7	778	12	ADJ41684	Adj41684 Plant cDN
25	587.5	14.4	1374	13	ADT48573	Adt48573 Bacterial
26	587.5	14.4	1518	6	ABL41880	AbL41880 Nucleotid
27	580	14.2	583	13	ACN56953	Acn56953 Cotton gy
28	562	13.8	870	6	ABK51964	Abk51964 Rice cDNA
29	562	13.8	870	8	ABX93200	Abx93200 cDNA enco
30	562	13.8	870	12	ADG47909	Adg47909 Rice Arab
31	552	13.5	659	12	ADJ41685	Adj41685 Plant cDN
32	548	13.4	644	12	ADJ42193	Adj42193 Plant cDN
33	539	13.2	1395	6	ABK73616	Abk73616 Bacillus
34	534	13.1	2127	12	ADM47936	Adm47936 Polynucle
35	533	13.1	1386	13	ADT48620	Adt48620 Bacterial
36	520.5	12.8	1826	3	AC42332	Ac42332 Arabidopo
37	513	12.6	575	13	ACN46405	Acn46405 Cotton pr
38	505	12.4	1422	13	ADT48346	Adt48346 Bacterial
39	499	12.2	536	13	ACN56998	Acn56998 Cotton gy
40	496	12.2	547	13	ACN48128	Acn48128 Cotton pr
41	496	12.2	1853	6	ABK51973	Abk51973 Soybean c
42	496	12.2	1853	8	ABX93209	Abx93209 cDNA enco
43	496	12.2	1853	12	ADG47927	Adg47927 Soybean B
44	494.5	12.1	2017	6	ABK51972	Abk51972 Rice cont
45	494.5	12.1	2017	8	ABX93208	Abx93208 cDNA enco

ALIGNMENTS

RESULT 1  
ADG47935  
ID ADG47935 standard; cDNA; 2777 BP.

XX AC ADG47935;

XX DT 11-MAR-2004 (first entry)

XX DE Corn Arabidopsis-like sugar transport protein cDNA #2.

XX KW Arabidopsis-like sugar transport protein;

XX KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;  
XX KW corn; plant; gene; ss.

XX OS Zea mays.

XX FH Key Location/Qualifiers

XX CDS 12..2414

XX FT /\*tag= a

XX FT /product= "Corn Arabidopsis-like sugar transport protein"

XX FT /note= "No start codon"

XX FT /partial

XX PN US2002199217-A1.

XX XX 26-DEC-2002.

XX PP 17-JAN-2002; 2002US-00051909.

XX

PR 24-APR-1998; 98US-0083044P.  
 PR 14-APR-1999; 99US-00291922.  
 PA (HELE/) HELENTJARIIS T G.  
 XX Helentjaris TG;  
 PI  
 XX WPI; 2004-040967/04.  
 DR P-PSDB; ADG47936.  
 DR  
 XX New isolated polynucleotide encoding a polypeptide having sugar transport  
 PT protein activity, for producing a transformed plant and for use as probes  
 PT in physical mapping.  
 XX  
 XX Claim 6; SEQ ID NO 31; 71pp; English.  
 PS  
 CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar  
 CC transport proteins and their corresponding nucleic acid sequences. The  
 CC sequences of the invention are useful to transform a cell. These are also  
 CC useful to produce a transgenic plant. Probes derived from sequences  
 CC encoding sugar transport protein may be used for physical mapping. The  
 CC present sequence is corn Arabidopsis-like sugar transport protein cDNA.  
 XX  
 XX Sequence 2777 BP; 667 A; 587 C; 701 G; 822 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0 Length: 2777  
 Score: 4075.00 Matches: 800  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-051-909-32 (1-800) x ADG47935 (1-2777)

QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArg 20  
 DB 12 ATTCGGAGCGCTCTTGGCTTGCAGTCCAAACGCGCTTCCACCCCTGATCTGGACCGGAGG 71  
 QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40  
 DB 72 GAGCGGCTCTTCGTCAGTGTCTTCTTGGCTGGGCGCTCTTCGCGCTGCTCGGTGT 131  
 QY 41 SerSerGluGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60  
 DB 132 TCITTCACAGGCGCGGTGACCTCGACGATATCTTGGAGACAAGATGTGGGGGCTGTT 191  
 QY 61 LeuValAlaIleValAlaSerIleGlyAenLeuLeuGlnGlyTrpAspAsnAlaThrIle 80  
 DB 192 CTGTGCGCCATAGTCGCTCCATCGCAATCTATTGACGGGTGGGACATGCCACCATC 251  
 QY 81 AlaAlaAlaValLeuTrpIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100  
 DB 252 GCAGCTGCTGTTCTGTATATAAAGAGGAATTCAAATGCAAAATGAGCCCACTGTGGAG 311  
 QY 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120  
 DB 312 GGACTAATTTGTCAATGTCTATATCGCGCCACCATCGTTACTACATTCCTCCGGGCA 371  
 QY 121 LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTrpPhe 140  
 DB 372 TTATCAGACTCGATTGGCGACGCCCTATGCTTATTTCTCTCAATTCGTACTTCTTC 431  
 QY 141 SerGlyLeuIleMetLeuTrpSerProAsnValTrpValLeuLeuAlaArgPheVal 160  
 DB 432 AGCGGCTCATCATGCTATGCTCTCCTAATGCTATGCTCTGCTGTGGCAGCTTCGTA 491  
 QY 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTrpIleSerGluIleAla 180  
 DB 492 GATGGATTGGTATTTGGCTTGGCTGTGACGCTTGGCTTGTACATTTTCAGAAATAGCC 551  
 QY 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200

DB 552 CCTTCGGAGATTAGAGTTTGTCTGAATACACTACCAATTCAGTGGATCAGGAGGAATG 611  
 QY 201 PheLeuSerTrpCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220  
 DB 612 TTCTTGTCTACTGTCATGCTGTTGGGATGTCCTGTGCCATCACCCGATGGAGAAAT 671  
 QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTrpLeu 240  
 DB 672 ATGCTTGGTGTCTCGGATACCTTCATTTCTTTTGGTTTGACAATATTTATCTT 731  
 QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysValLeu 260  
 DB 732 CCTGAATCTCCAGATGGCTCTTACCAAGATCGATCGCAGAGGCAAAAGAGTGTG 791  
 QY 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280  
 DB 792 CAAAGTTTACGGGGGAAAGACGATGTCTCAGGTGAATTTGCTCTTCTTCAAGGGTGG 851  
 QY 281 GluValGlyGlyAspThrSerIleGluGluTrpIleIleGlyProAlaThrGluAla 300  
 DB 852 GAGGTTGGAGGAGACACTTCCATTTGAAGAGTACATCATTTGGACCTGCCACCGAGCAGCC 911  
 QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTrpGlyProGluGluGly 320  
 DB 912 GATGATCTTGTACTGACGGTGATTAAGAAACAAATCACACTTTATGGGCTTGAAGAAGGC 971  
 QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340  
 DB 972 CAGTCATGATGCTCTGACCTTCTAAGGGACCCATCATGCTTGGAAAGTGTCTTCTCT 1031  
 QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360  
 DB 1032 GCATCTCGTCATGGGAGCATGTTGAACAGAGTGTACCCCTTATGGATCCGATCTGACA 1091  
 QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380  
 DB 1092 CTTTTTGGTAGTGTCTCATGAGATATGCTCTCAAGCTGGAGGAAGTATGAGGAGCACA 1151  
 QY 381 PheProAspPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGlnTrp 400  
 DB 1152 TTTTCAAACCTTTGGAAAGTATGTTCAAGTGTCAAGATCAGCATGCCAAATATGACGATGG 1211  
 QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTrpAlaSerAspGlyAlaGlyAsp 420  
 DB 1212 GATGAAGAGAAATCTTCATAGGGATGACGAGGAGTACGCACTCTGATGGTGCAGAGGTGAC 1271  
 QY 421 TrpGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440  
 DB 1272 TATGAGACAATCTCCATAGCCCATTTGCTGTCCAGGACGAGCAACAGGTGCGGAAGGGAAG 1331  
 QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460  
 DB 1332 GACATTTGTCACCATGTCACCGTGGAGTGTCTTTGAGCATGAGAAGGCAAAACCCCTCTTA 1391  
 QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480  
 DB 1392 GGGAGGGTGGAGATGGTGTGACGACACTGATATCGGTGGGGGATGGCAGCTTGTCTGG 1451  
 QY 481 LysTrpSerGlnLysGluGlyGluAsnGlyArgLysGlyGlyGlyPheLysArgValTrp 500  
 DB 1452 AAATGGTGCAGAGAAGGAGTGAATGTGTAAGAGAGGTGGTTCCTTCAAAAGAGTCTAC 1511  
 QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520  
 DB 1512 TTGCACCAAGAGGAGTTCCTGGCTCAAGAAGGGGCTCAATTTGTTTCACTTCCCGGTGGT 1571  
 QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540  
 DB 1572 GGGCATGTTTTTGGAGGTAGTGTGATTTGTACATGCTGCTTGTAGTAAGTCAAGTCA 1631  
 QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560  
 DB 1632 CTTTTTCTCAAAGGGTCTTGTCTGAACACCGCATGTGATGCTGCCATGTTCCCATCTCT 1691

```
QY 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAla 580
|
|
|
Db 1692 GAGGTAGCTGCCAAGGTTCCACGTTGGAAAGATTGTTGTTGAACCTCGAGTCGCGTCC 1751
|
|
|
QY 581 LeuLeuValGlyValGlyLeuGlnLeuGlnPheAlaGlyLeuAenGlyValLeu 600
|
|
|
Db 1752 CTGTTAGTCGGTGTGGAAATTCAGATCCCTCAACAGTTTCTGGAAATAAACGGTGTCTG 1811
|
|
|
QY 601 TyrTyrThrProGlnLeuLeuGluGlnAlaGlyValAlaValLeuSerLysPheGly 620
|
|
|
Db 1812 TACTATACCCACAAATCTTGGACAGCTGGTGGCAGTTATCTTTCCAAATTTGGT 1871
|
|
|
QY 621 LeuSerSerAlaSerAlaSerLeuLeuLeuSerSerLeuThrThrLeuLeuMetLeuPro 640
|
|
|
Db 1872 CTCAGCTCGGCATCAGCATCCATCTTGTATCAGTGTCTCTCACTACCTTACTTAATGCTTCT 1931
|
|
|
QY 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGly 660
|
|
|
Db 1932 TGCATTTGGCTTTGCCATGCTGCTTATGGATCTTTCGGAAAGAGTTTTCGTCGTAGGC 1991
|
|
|
QY 661 ThrIleProIleLeuLeuAlaSerLeuValLeuValLeuValSerAsnLeuIleAspLeu 680
|
|
|
Db 1992 ACAATTCCAATCTTGATAGCATCTCTAGTTATCTCTGTTGTGCCAATCTAATTTGATTTG 2051
|
|
|
QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
|
|
|
Db 2052 GGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATATCGTCTACTTCTGCTGCTTC 2111
|
|
|
QY 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
|
|
|
Db 2112 GTTATGGGATTTGGTCCATCCCAACATTTATGTGACAGATCTTTCACACAGGGTT 2171
|
|
|
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
|
|
|
Db 2172 CGTGGCTCTGTATTGCCATTTGCTGCTTTACATCTTGGATCGAGATATCATCGTCACC 2231
|
|
|
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
|
|
|
Db 2232 TACAGCCTTCTCTGTGCTGCTGAATGCTATTGGACTGCGCGGTGTTTTTCAGCATATATGCA 2291
|
|
|
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
|
|
|
Db 2292 GTCGTATGCTTGATTTCTTTGTTGTTGCTCTTCTTAAGGTCCCTGAGACAAAGGGGATG 2351
|
|
|
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
|
|
|
Db 2352 CCCCTTGAGGTTATTACCGAATTTCTTTCGAGTTGGTGGAGCAAGCGGCTGCAAAAGCC 2411
|
|
|
RESULT 2
ADM47932
ID ADM47932 standard; DNA; 2908 BP.
XX
AC ADM47932;
XX
AC 1
XX
DT 03-JUN-2004 (first entry)
XX
DE Polynucleotide sequence #350 useful in producing transgenic plants.
XX
KW Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;
KW osmotic stress; sugar transport; cell cycle pathway; plant height;
KW carbohydrate transport; crop productivity; plant growth;
KW stress resistance; disease resistance; insect resistance; heat tolerance;
KW nitrogen assimilation; water stress tolerance;
KW photosynthetic carbon fixation; virus resistance; gene therapy; gene; ds.
XX
OS Zea mays.
XX
PN US2003233670-A1.
XX
PN 18-DEC-2003.
PD
XX
XX 04-DEC-2002; 2002US-00310154.
PF
```

```
XX
PR 04-DEC-2001; 2001US-0337358P.
XX
PA (EDGE/) EDGERTON M D.
PA (CHOM/) CHOMET P S.
PA (LACC/) LACCETTI L B.
XX
PI Edgerton MD, Chomet PS, Laccetti LB;
XX
XX WPI; 2004-061374/06.
DR DR P-PSDB; ADM48300.
XX
PT New polynucleotide, useful for manipulating plant protein quality,
PT improving plant growth, yield and crop productivity or grain composition
PT or producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 350; 144pp; English.
XX
XX The present invention relates to polynucleotide sequences, and the
XX proteins they encode. The sequences are isolated from a variety of
XX organisms such as plants (e.g. maize, rice, sorghum, thale cress,
XX soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The
XX polynucleotide and polypeptide sequences of the invention are useful in
XX the production of transgenic plants that have improved properties. Also
XX disclosed are methods of producing fertile transgenic plants, preferably
XX maize, with desired phenotypes. The polynucleotide and polypeptide
XX sequences are useful for improving plants by providing protection against
XX osmotic stress, improving altering sugar transport and/or metabolism,
XX modifying the cell cycle pathway, reducing plant height, modifying
XX carbohydrate transport, improving crop productivity, improving plant
XX growth and stress resistance, improving cold or heat tolerance, improving
XX insect resistance, improving stalk strength, improving water stress tolerance,
XX assimilating, improving photosynthetic carbon fixation, improving biotic and abiotic
XX stress resistance, improving resistance to oxidative stress, providing
XX increased vigour, reducing senescence, and conferring virus resistance.
XX The present sequence represents a polynucleotide sequence of the
XX invention. Note: the sequence data for this patent is not provided in the
XX printed specification but is obtained in electronic format from the USPTO
XX website at seqdata.uspto.gov.
XX
SQ Sequence 2908 BP; 669 A; 640 C; 747 G; 852 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 2908
Score: 4059.00 Matches: 797
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 12 Gaps: 0
XX
US-10-051-909-32 (1-800) x ADM47932 (1-2908)
QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArg 20
|
|
|
Db 166 ATTCGAGCGGCTCTTGGCTTGCGATCCAGACGCCCTTCACCCCTGATCTGGACCGGAGG 225
|
|
|
QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
|
|
|
Db 226 GAGCGACTCTTCGGTCAGTTGTTCTTGTCTTGGCTGGGCCCTCTTCGCCCTGCTTCGTGT 285
|
|
|
QY 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
|
|
|
Db 286 TCTTCAAGAGCGGCTGACCTCGGACGATATCTTGGAGGACAAGATGTCGGGGGCTGTT 345
|
|
|
QY 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAlaThrIle 80
|
|
|
Db 346 CTGTGCGCATAGTCGCCCTCCATTCGCAATCTATTTCAGGGGTGGGACAATGCCCATC 405
|
|
|
QY 81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
|
|
|
Db 406 GCAGCTGCTGTTCTGTATATTAAGAGGAATTCATTTGCATAATGAGCCCACTGTGGAG 465
|
|
|
```

101 GlyLeuValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120  
111  
466 GGACTAAATTTGTGTCACATGTCACATGTCGGGCCACCATGCTTACTACATCTCCGGGCCA 525  
121 LeuSerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140  
131  
526 TTATCAGACTCGATTGGCCGAGCCCTATGCTTATTTCTCTTCAATTCGTACTTCTTC 585  
141 SerGlyLeuIleMetLeuTrpSerProMetValTyrValLeuLeuAlaArgPheVal 160  
151  
586 AGCGGCTCATCATGCTATGGTCTCTAATGCTATGCTGCTGCTGGCCAGCTTGGTA 645  
161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180  
191  
646 GATGGATTGGTATTTGGCTGTGCTGCACGCTGTGCTTGTATCATTTTCAGAAATAGCC 705  
181 ProSerGluIleArgGlyLeuLeuMetSerLeuProGlnPheSerGlySerGlyMet 200  
191  
706 CTTTCGAGATTAGAGTTTGTGGAATACACTACACCAATTCAGTGGATCAGGAGGAATG 765  
201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220  
211  
766 TTCTTGTCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825  
221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240  
231  
826 ATGCTTGGTGTCTCCGATACCTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 885  
241 ProGluSerProArgTrpLeuValSerIleGlyArgMetAlaGluAlaLeuValLeu 260  
251  
886 CTTGAATCTCCAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 945  
261 GlnLeuLeuArgGlyIleAspValSerGlyGluLeuSerLeuLeuLeuGlyLeu 280  
291  
946 CAAAGTTACGGGGGAAAGACATGCTCAGGTGAATTTGCTCTTCTTCTTCTTCTTCTT 1005  
281 GluValGlyGlyAspThrSerIleGluIleGlyIleGlyProAlaThrGluAlaAla 300  
311  
1006 GAGTTTGGAGGAGACATCTCCATTGAAGAGTACATCTGACCTGCCACCGAGGAGCC 1065  
301 AspAspLeuValThrAspGlyAspIleGluIleThrLeuTyrGlyProGluGluGly 320  
331  
1066 GATGATCTTTGTTACTGACGGTGATGAAGAAACAAATCACATTTATGGGCTGAAGAGGC 1125  
321 GlnSerTrpIleAlaArgProSerIleGlyProIleMetLeuGlySerValLeuSerLeu 340  
351  
1126 CAGTCATGATTTGCTGACCTTCCAGGGACCCAGCATGCTTGGAGTGTGCTTCTTCTT 1185  
341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360  
371  
1186 GCACTCTGTCATGGGAGCATGCTGAACAGAGTGTACCCCTTATGATCCGATTTGTGACA 1245  
361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380  
391  
1246 CTTTTTGGTAGTGTCCATGAGATATGCTCAAGCTGGAGGAAGTATGAGGAGCAATTG 1305  
381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLeuAsnGluGlnTrp 400  
411  
1306 TTTTCCAACTTTGGAGATGTTCAAGTGTCAAGATCAGCATGCCCAAAATAGAGAGTG 1365  
401 AspGluGluAsnLeuHisArgAspAspGluIleTyrAlaSerAspGlyAlaGlyAsp 420  
431  
1366 GATGAAGAGAATCTTCATAGGATGACGAGGAGTACGATCTGATGGTGCAGGAGGTGAC 1425  
421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440  
451  
1426 TATGAGGACAAATCTCCATAGCCATTTGCTGTCAGGAGGCAACAGGTGCGGAGGGAAG 1485  
441 AspileValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460  
471  
1486 GACATTTGCACCATGCTCACCCTGGAGTGTCTTTGAGCATGAGAAAGCCCTCTTA 1545  
461 GlyGluGlyGlyAspGlyValSerSerThrAspileGlyGlyTyrTrpGlnLeuAlaTrp 480

1546 GGGGAGGGTGGAGATGGTGTGAGCAGCATGATATCGGTGGGGGATGCGACTGCTTGG 1605  
481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGlyGlyPheLysArgValTyr 500  
1606 AAATGGTCAGAGAAGAGGTGAGAAATGGTAGAAAGAGGTGGTGTTCAAAGAGAGTCTAC 1665  
501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520  
1666 TTGCACCAAGAGGGATTCCTGGCTCAAGAAGGGCTCAATTTGTTTCACTTCCCGTGGT 1725  
521 GlyAspValPheGluGlySerCluPheValHisAlaAlaLeuValSerGlnSerAla 540  
1726 GCGATGTTCTTGAGGGTAGTGTGTCATGCTGCTGCTTGTAGTAGTACGTACGCA 1785  
541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560  
1786 CTTTCTCAAGGGCTTGTGTCACGATGTCAGATGCTGCCATGTTCCATCTC 1845  
561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAla 580  
1846 GAGGTAGCTGCCAAAGGTTACGTTGGAAAGATTTGTTGAACCTGGAGTGGGGTGC 1905  
581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600  
1906 CTGTTAGTCGGTGTGGAATTCCTTCAACAGTTTGTGCGAATTAACCGGTGTTCTG 1965  
601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620  
1966 TACTATACCCCAAAATCTTGAGCAAGCTGGTGTGGCAGTTATTTCTTTCCAAATTTGGT 2025  
621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640  
2026 CTGAGTCTGGCATCAGCATCCATCTTGATCAGTTCTCTCACCTACTACTAATGCTTCT 2085  
641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGly 660  
2086 TGATTTGGCTTTGCCATGCTGCTTATGATCTTTCCGAGAGAGGTTTTTGTCTGTAGGC 2145  
661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680  
2146 ACAATTCCAATCTTGATAGCATCTCTAGTTATCTGTTGTTGTCACATCTAATGATTTG 2205  
681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700  
2206 GGTACACTAGCCCATGCTTGTCTCCACCGTCAGTGTATGCTACTTCTGTGCTTTC 2265  
701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720  
2266 GTTATGGGATTTGGTCCCATCCCAACATTTTATGTCAGAGATCTTTCCAAACAGGGTT 2325  
721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspileIleValThr 740  
2326 CGTGGCTCTGTATTGGCAATTTGCTTTTACATTTCTGATCGAGATATCATCTCACC 2385  
741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760  
2386 TACAGCTTCTCTGTGATGCTGATGCTATGACTGGCGGGTGTTCAGCATATATGCA 2445  
761 ValValCysLeuIleSerPheValPheLeuLysValProGluThrLysGlyMet 780  
2446 GTCGTAGCTGATTTCTTTGTTGCTTCTTAACTCCCTGAGACAAAGGGGATG 2505  
781 ProLeuGluValIleThrGluPheAlaValGlyAlaLysGlnAlaAlaLysAla 800  
2506 CCCCTTTGAGGTTATTACCGAATTTCTTTGAGTTGGTGGAGAGCGGCTGCAAAAGCC 2565

RESULT 3  
ABK51962  
ID ABK51962 standard; cDNA; 2824 BP.  
XX  
AC ABK51962;  
XX





```

Db 1498 CGGGGATGGCAGCTTGGCTTGGAAATGGTCAGAGAAGGAGCTGAGAATGGTAGAAGGAA 1557
Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgAtgGlySer 513
Db 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTGGCTCAAGAAGGGGCTCA 1617
Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 1618 ATGTGTTTCACTTCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTGTACATGCTGCT 1677
Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 1678 GCTTTAGTAGAAGTCAGTCAGCACTTCTTCAAGGGTCTTGCTGAACCAACGATGTCAGAT 1737
Qy 554 AlaAlaMetValHisProSerGluValAlaLysGlySerArgTyrLysAspLeuPhe 573
Db 1738 GCTGCGCATGGTTTCAACCATCTAGATGCTGCGAAGGTTTCACTTGAAGATTTGTTT 1797
Qy 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db 1798 GAACCTGGAGTGAGCGTCCCTGTGTAGTCGGTGTGGAATTCAGATCCTTTCAACAGTTT 1857
Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 1858 GCTGGAATAAACCAGTGTCTGTACTATACCCACAAATTTCTTGAGCAAGCTGGTGTGGCA 1917
Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTTATCTTTCCAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC 1977
Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACTTACTAATGCTTCTTGCAATGGCTTTCGCAATGCTGTGATGATCTTTCGGA 2037
Qy 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 2038 AGAAGGTTTTTCTGCTGAGGCAATTCCAATCTTGATAGCATCTCTAGTTATCTCTGTT 2097
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCAATCTAATTTGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157
Qy 694 IleValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATCGTCTACTTCTGCTGCTTCTGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217
Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr 733
Db 2218 GAGATCTTTCCAAACAGGGTTCGTGGCTCTGTATTGCAATTTGTGCTTTTACATTTCTGG 2277
Qy 734 IleGlyAspIleIleValTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCGGAGATATCATGTCACCTACAGCTTCTGTGATGCTGTAATGCTATTGGACTGGCG 2337
Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLys 773
Db 2338 GGTGTTTTTCAAGATATATGACGTGATGCTTGATTTCTTGTGTTGCTTCTTAA 2397
Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTGACAAAGGAGATGCCCTTGAGTTATTACCGAATTTCTTGGCAGTTGGTGGC 2457
Qy 794 LysGlnAlaAlaLysAla 800
Db 2458 AAGCAAGCGCTGCAAAAGCC 2478

```

RESULT 4

ABX93198

ID ABX93198 standard; cDNA; 2824 BP.

XX

AC ABX93198;

XX

DT 29-MAY-2003 (first entry)

```

XX DE cDNA encoding corn sugar transport protein #1.
XX KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX OS Zea mays.
XX PN US2002178468-A1.
XX PD 28-NOV-2002.
XX PF 17-JAN-2002; 2002US-00051902.
XX PR 24-APR-1998; 98US-0083044P.
XX PR 14-APR-1999; 99US-00291922.
XX PA (ALLEN/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2003-340957/32.
XX P-PSDB; ABU08326.
XX PT Novel plant sugar transport proteins and nucleic acid encoding the
XX PT protein useful for producing transgenic plants having altered levels of
XX PT sugar transport protein.
XX PS Claim 2; Page 13-15; 56pp; English.
XX CC The present invention relates to the isolation of Arabidopsis thaliana-
XX CC like or Beta vulgaris-like sugar transport proteins, and the
XX CC polynucleotide sequences encoding them. The plant sugar transport
XX CC proteins of the invention have been isolated from corn, rice, soybean,
XX CC and wheat. The polypeptides of the invention may be used for altering the
XX CC level of expression of a sugar transport protein in a host cell, by
XX CC transforming a host cell with a chimeric construct encoding all, or a
XX CC portion of the sugar transport protein, in sense or antisense
XX CC orientation. Particularly, the polypeptides may provide a means to
XX CC control carbohydrate transport and distribution in plants. ABX93198-
XX CC ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like
XX CC sugar transport proteins
XX SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.: 4.45e-298 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: Gaps: 2

US-10-051-909-32 (1-800) x ABX93198 (1-2824)
Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTyr 75
Db 238 ATGGGGGGCGCGCTGATGTCGCCATCGGCTCTATCGCACTTCTCGAGGCTGG 297
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 298 GACAAATGGCACAATTTGCTGGAGCCCTCTGTACATAAAGAAGGAATTCACCTGCAGAGC 357
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTCTGATCGAAGGCTCATCGTCGCCATGTTCCTCATTTGGGGCAACAGTATCACA 417
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135

```

418	Db	ACATCTCCGGGGCCCAAGGGCTGAC	CTGCGTTGGTGTAGGAGGCCCACTGCTGTGGCTCGGCTCGGCT	477
136	Qy	IleLeuTyrPhePheSerGlyLeuIle	MetLeuTrpSerProAenValTyrValLeuLeu	155
478	Db	GTCTCTACTTCTGTCAGTGGGCTGGT	GATGCTTTGGGGCCCAATTGGTGTACATCTTGGCTC	537
156	Qy	LeuAlaArgPheValAspGlyPheGlyIle	GlyLeuAlaValThrLeuValProLeuTyr	175
538	Db	CTCCAAAGGCTCATGTGATGGGTTCCGGT	ATCGGTTTGGCGGTCAACATGTGTCTCTCTCTAC	597
176	Qy	IleSerGluIleAlaProSerGluIle	Arg---GlyLeuLeuAenThrLeuProGlnPhe	194
598	Db	ATCTCCGAAACTGCACCGCACAGANAT	CTTGGGGCTGNTNGAACAGTGTCCGCGAGTTC	657
195	Qy	SerGly---SerGlyGlyMetPheLeuSer	TyrCysMetValPheGlyMetSerLeuSer	213
658	Db	ATTGGGGTCAGNGAGGGATGTTCTCTCT	CTACTGCATGGTGTGGATGTCCCTCATG	717
214	Qy	ProSerProAspTrpArgIleMetLeuGly	ValLeuAlaIleProSerLeuPhePhe	233
718	Db	CCCAAACTGATTGGAGGCTCATGCTTGGAG	TTCGTGCGATCCGTCACATTATNTACTTT	777
234	Qy	GlyLeuThrIlePheTyrLeuProGluSer	ProArgTrpLeuValSerIysGlyArgMet	253
778	Db	GGACTGACTGTCTTCTACTTGCCTGGAAT	CACCAAGGTGGCTGTGTNAGCAAGAGAGATG	837
254	Qy	AlaGluAlaIysIysValLeuGlnIysLeu	ArgIysAspValSerGlyGluLeu	273
838	Db	CGGAGGGGAAGAGATGNTGCCAAGGCT	CGGGGAAGAGAGATGTCTCANGGGAGANG	897
274	Qy	SerLeuLeuLeuGluGlyLeuGluValGly	GlyAspThrSerIleGluIuTyrIleIle	293
898	Db	GCTCTTCTAGTTGAAGCTTTGGGGCTCGG	TAAAGATACACGTATTTNAGAGTACATCATT	957
294	Qy	GlyProAlaThrGluAlaAlaAspAspLeu	ValThrAspGlyAspIysGluGlnIleThr	313
958	Db	GGACCTGCCACCGAGGACGCGATGTCTGT	GAATCAGCGGTGATAAGGAACAAATCACAC	1017
314	Qy	LeuTyrGlyProGluGluGlyGlnSerTrp	IleAlaArgProSerIysGlyProIleMet	333
1018	Db	CTTTATGGGCTGAAAGAAGGCCAGTCAATG	ATGTCTGCACCTTCTAAGGAGCCCATCATG	1077
334	Qy	LeuGlySerValLeuSerLeuAlaSerArg	HisGlySerMetValAenGlnSerValPro	353
1078	Db	CTTGGAAGTGTGCTTTCTCTTGCACTCGT	CATGGGAGCATGGTGAACACAGAGTGTACCC	1137
354	Qy	LeuMetAspProIleValThrLeuPheGly	SerValHisGluAenMetProGlnAlaGly	373
1138	Db	CTTATGATCCGATGTGTGACACTTTTGGT	AGTGTCCATGAGAATATGSCCTCAAGCTGGA	1197
374	Qy	GlySerMetArgSerThrLeuPheProAsp	NPheGlySerMetPheSerValThrAspGln	393
1198	Db	GGAAAGTATGAGAGCACATTGTTTCCAAAC	TTTTGGAAAGTATGTTTCACTGTGCACAGATCAG	1257
394	Qy	HisAlaIysAenGluGlnTrpAspGluGlu	AenLeuHisArgAspAspGluTyrAla	413
1258	Db	CATCCCAAAATGAGCAGTGGGATGAAGAA	ATCTTCTATAGGATGACGAGGAGTACGCA	1317
414	Qy	SerAspGlyAlaGlyGlyAspTyrGluAsp	AenLeuHisSerProLeuLeuSerArgGln	433
1318	Db	TCTGATGGTCAGAGGATGACTATGAGGAC	AAATCTCCATAGCCCATTTGCTGTCCAGCGAG	1377
434	Qy	AlaThrGlyAlaGluGlyIysAspIleVal	HisHisGlyHisArgGlySerAlaLeuSer	453
1378	Db	GCAACAGGTGGGAAGGAAGGACATTGTG	CACCATGGTCACCGTGAAGTGTCTTGAGC	1437
454	Qy	MetArgArgGlnThrLeuLeuGlyGluGly	GlyAspGlyValSerSerThrAspIleGly	473
1438	Db	ATGAGAGGCCAAACCCCTTAGGGGAGGG	TGGAGATGGTGTGAGCACCACTGATATCGGT	1497
474	Qy	GlyGlyTrpGlnLeuAlaTrpLysTrpSer	GlnIysGluGlyIysGluAenGlyArgLysGlu	493
1498	Db	GGGGGATGGGACTTGTCTGGAATGTGT	CAGAGAAGGAAGTGTGAGATGTGTAGAAAGGAA	1557

Qy	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513
Db	1558	GGTGGTTTCAAAAGAGTCTACTTTCGACCAAGAGGGAGTTCCTGGCTCAAGAAAGGGGCTCA	1617
Qy	514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1618	ATTGTGTTTCACTTCCCGGTGGTGGCATGTTCTTTGAGGGTAGTAGTGTGTGTACATGCTGCT	1677
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1678	GCITTAGTAGTCAGTCAGGACATTTTCTCAAGGGTCTTCTGCTGNACCACCGCATGTGCAGAT	1737
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe	573
Db	1738	GCTGCCATGGTTCACCCATCTGAGTAGCTGCCAAAGGTTTCAAGTTCGCAAGATTTCTGTT	1797
Qy	574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe	593
Db	1798	GAACCTCGAGTCAGGCGGTGCCCTGTTAGTCGGTGTTGGAAATTCAGATCCCTTCAACAGTGT	1857
Qy	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613
Db	1858	GCTGGAAATAACGGTGTTCTGTACTATACCCCAAAATTTCTTGAGCAAGCTGGTGTGGCA	1917
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu	633
Db	1918	GTTATTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTCTCTCTC	1977
Qy	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1978	ACTACCTTACTAAATGCTTCTTCGATTGGCTTTTGCCATGCTGCTTATGSGATCTTTCCGGA	2037
Qy	654	ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	2038	AGAAGGTTTTGCTGTAGGCACAAATTCCAATCTTGATAGCATCTCTAGTTATCTCTGGTT	2097
Qy	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	2098	GTGTCCAAATCTAATTGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT	2157
Qy	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2158	ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATGCCCAACATTTTATGTGCA	2217
Qy	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp	733
Db	2218	GAGATCTTCCAACCGGGTTCGTGGGCTCTGTATTGCCATTTGTGCTTTACATCTCG	2277
Qy	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2278	ATCGAGATATCATCGTCACCTACAGCCTTCTCTGTGTGCTGAATGCTATTGGAGTCGCG	2337
Qy	754	GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys	773
Db	2338	GGTGTTTTCAGCATATATGCAATGTCATGCTGATTTCCCTTTGTGTTGCTCTCTCTTAAG	2397
Qy	774	ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla	793
Db	2398	GTCCCTGAGACAAAGGGGATGCCCTTGAGGTATTATCCGAATTTCTTTGCAGTTGTGGCG	2457
Qy	794	LysGlnAlaAlaLysAla 800	
Db	2458	AAGCAAGCGCTGCAAAAGCC 2478	
RESULT 5			
ID	ADG47905		
XX	ADG47905 standard; cDNA; 2824 BP.		
AC	ADG47905;		
XX	11-MAR-2004 (first entry)		
DT	XX		



Db	1318	TCTGATGGTGCAGAGGAGTACTATAGGACAATCTCCATAGCCCATGCTGTCCAGGCAG	1377
Qy	434	AlaThrGlyAlaGluGlyIysAspIleValHisGlyHisArgGlySerAlaLeuSer	453
Db	1378	GCAACAGGTGGGAGGGAAGGACATTGTGCACCATGGTCACCGTGGAGTGTCTTGAGC	1437
Qy	454	MetArgArgGlnThrLeuLeuGlyGluGlyIysAspGlyValSerSerThrAspIleGly	473
Db	1438	ATGAGAAGGCAAGACCTCTTAGGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGT	1497
Qy	474	GlyGlyTyrGlnLeuAlaTyrIysTyrSerGluLysGluGlyAlaGlyArgLysGlu	493
Db	1498	GGGGGATGGCAGCTTCTTGGAATAGTGTGAGAGGAGAGTGTGAGATGTGTAGAAAGGA	1557
Qy	494	GlyGlyPheIysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513
Db	1558	GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCCTGGCTCAAGAGGGCTCA	1617
Qy	514	IleValSerLeuProGlyGlyIysAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1618	ATTGTTTCACTTCCCGTGGTGGCATGTTCTTGAGGGTAGTGAGTGTGTACATGCTGCT	1677
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1678	GCTTTAGTAGTCAGTCAGCACCTTTCTCAAGGGTCTTGCTGAACCGCATGTGCAGAT	1737
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe	573
Db	1738	GCTGCCATGGTTCACCACTGAGGTAGTGTCCCAAGGTTTCAGTTGGAAAGATTGTGTT	1797
Qy	574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnInPhe	593
Db	1798	GAACCTGGAGTGGGGTGGCTGTGTAGTCGGTGTGGATTTCAGATCCTTCAACAGTTT	1857
Qy	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613
Db	1858	GCTGGAATAAACGGTGTCTGTACTATACCCCAAAATTCTTGAGCAAGCTGTGTGGCA	1917
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu	633
Db	1918	GTTATTTCTTCCAAATTTGGTCTCAGCTGGCATCAGCATCTCTTGATCAGTTCTTC	1977
Qy	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1978	ACTACCTTACTAATGCTTCTTGCATTTGGCTTTGGCATGCTGTATGATCTTCCGGA	2037
Qy	654	ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	2038	AGAAAGGTTTTTGTGTGTAGGCACAAATTCCAATCTTGATAGCATCTCTAGTTATCT	2097
Qy	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	2098	GTGTCCAACTAATGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT	2157
Qy	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2158	ATCGTCTACTTCTGCTGCTTGTATGGGATTTGGTCCCATCCCCAACCAATTTATGCA	2217
Qy	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr	733
Db	2218	GAGATCTTTCCAAACCGGGTTCGTGGCCCTGTATTGGCCATTTGTGCTTTTACATTC	2277
Qy	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2278	ATCGGAGAT	2337
Qy	754	GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys	773
Db	2338	GGTGTGTTTACAGAT	2397
Qy	774	ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla	793

Db	2398	GTCCCTGACACAAAGGGATGCCCTTGAGGTATTATACGAATCTTTGCAGTTGGTGGC	2457
Qy	794	LysGlnAlaAlaAlaLysAla	800
Db	2458	AAGCAAGCGCTGCAAAAGCC	2478
RESULT 6			
ABK51965			
ID	ABK51965	standard; cDNA; 2601 BP.	
XX	AC	ABK51965;	
XX	DT	27-AUG-2002 (first entry)	
XX	DE	Soybean contig encoding A. thaliana-like sugar transport protein.	
XX	KW	Soybean; Arabidopsis thaliana-like sugar transport protein;	
XX	KW	carbohydrate transport; grain filling; annual field crop; plant; geno;	
XX	OS	88.	
XX	XX	Glycine max.	
XX	XX	Location/Qualifiers	
XX	XX	175..2388	
XX	XX	/tag= a	
XX	XX	/product= "Soybean Arabidopsis thaliana-like sugar	
XX	XX	transport protein"	
XX	XX	US6383776-B1.	
XX	XX	07-MAY-2002.	
XX	XX	14-APR-1999; 99US-00291922.	
XX	XX	24-APR-1998; 98US-0083044P.	
XX	XX	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX	XX	Allen SM, Hitz WD, Kinney AJ, Tingey SV;	
XX	XX	WPI; 2002-453364/48.	
XX	XX	P-PSDB; AAU97204.	
XX	XX	New nucleic acid encoding plant sugar-transport proteins, useful for	
XX	XX	preparing transgenic plants with altered carbohydrate distribution.	
XX	XX	Claim 3; Col 35-38; 54pp; English.	
XX	XX	The present invention relates to the isolation of plant polynucleotide	
XX	XX	sequences encoding an Arabidopsis thaliana-like sugar transport protein	
XX	XX	or Beta vulgaris-like sugar transport protein. The polynucleotide	
XX	XX	sequences are useful for altering the level of sugar transport proteins	
XX	XX	in plants, i.e. for control of carbohydrate transport and distribution in	
XX	XX	plant cells, e.g. during grain filling of annual field crops (e.g. corn,	
XX	XX	rice, soybeans, and wheat), and, for studying carbohydrate flows and	
XX	XX	sugar transport. The polynucleotide sequences can also be used to isolate	
XX	XX	cDNA sequences and genes that encode homologues of the new proteins. The	
XX	XX	present sequence representing a contig assembled from various soybean	
XX	XX	cDNA clones encodes an Arabidopsis thaliana-like sugar transport protein	
XX	XX	Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;	
Alignment Scores:			
Pred. No.:	3..76e-224	Length:	2601
Score:	2674.00	Matches:	522
Percent Similarity:	82.20%	Conservative:	92
Best Local Similarity:	69.88%	Mismatches:	115
Query Match:	65.62%	Indels:	18
DB:	6	Gaps:	8
US-10-051-909-32 (1-800) x ABK51965 (1-2601)			
Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr	75

175 ATGAAAGTGGCGTCTTGTGCTATTCGCGCTTCCATTGGTAATTTCTCCAAGGATGG 234  
QY  
76 AspAsnAlaThrIleAlaAlaValLeuTyrlleLysGlySerGlnLeuGlnAsn 95  
Db  
235 GATAATGCTACCATCGCGGGCTAATGTTACATTAAGAAAGACCTTGTCTTGGGAACA 294  
QY  
96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Db  
295 -----ACTATGAAAGGCTTGTGTGGGCATGTCCTGATTGGAGCAACGGTAATCACC 348  
QY  
116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
Db  
349 ACATGCTCTGGTCTTATAGCGATTTGGCTCGGTGGCGGCCATGATGATNAATCACT 408  
QY  
136 IleLeuTyrllePhePheSerGlyLeuIleMetLeuThrProAsnValTyrlleValLeu 155  
Db  
409 GTGCTCTATTTCTTGGGTGTTTGGTGAATTTGGGATTTGGCTTGTGACTTGTTCCTCCGGCTAT 468  
QY  
156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrl 175  
Db  
469 TTGGCGAGGCTACTTGTATGGATTTGGGATTTGGCTTGTGACTTGTTCCTCCGGCTAT 528  
QY  
176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
Db  
529 ATATCTGAAACGGCGCTCGAAATAAGGGGTGCTTGAATACGCTTCTCAGTTCACT 588  
QY  
196 GlySerGlyGlyMetPheLeuSerTyrlleCysMetValPheGlyMetSerLeuSerProSer 215  
Db  
589 GACTCTGGAGGAATGTTTGTCTGCTACTGTATGTTTGTGGCATGTCAATGAGTCCGCG 648  
QY  
216 ProAspThrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
Db  
649 CTAAGTCTGAGCTCATGCTGGGTCTGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 708  
QY  
236 ThrIlePheTyrlleProGlnSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
Db  
709 ACCATTTTCTTCCCGAGTCTCCTCGGTGGCTGGTGCAGAAAGAGAGATGCTCGAG 768  
QY  
256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
Db  
769 GCTAAGAAAGGTCTCCAAAGATTTGGCGAAGAGGAGGATGTGTGAGCGAGATGGCATTT 828  
QY  
276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrlleIleGlyPro 295  
Db  
829 CTGGTTGAAGGTCTCGGATTTGGGGTGTATACATCTATCGAAGAGTACATAATTTGGCCCT 888  
QY  
296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGlyGluGlnIleThrLeuTyrl 315  
Db  
889 GTGACGATGTGGTGTGATGATCAACATCAACAGAGAGAGATAAAATTCGATTATAT 948  
QY  
316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
Db  
949 GGATCCCAAGCGGCTTTCTTGGTTATCAAACTGTCTACTGACAGAGTCTATTTGGC 1008  
QY  
336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
Db  
1009 -----CTTGGCTCACACATGGAGAGCATCATCAACAAAGCATGCCCCCTCATG 1056  
QY  
356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373  
Db  
1057 GATCTCTGGTGACACTGTTGGTAGCATTCATGAGAGGCTCCCGGAGAGAGAGCAAGA 1116  
QY  
374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
Db  
1117 GGAAGCATGCGAAGCACCTGTGTTCCAAATTTTGAAGCATGTTTCAGCATGCTGTGAGCCG 1176  
QY  
394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGluTyrlleAla 413  
Db  
1177 CATGCTAAATTTGAACATGGATGAGCAAGAAAGCTTACAAAGGAAACCTGAGGAGTACAT 1236  
QY  
414 SerAspGlyAlaGlyGlyAspTyrlleGluAspLeuHisSerProLeuLeuSerArgGln 433  
Db

1237 TCAGATGCAACCGCTGGGAGCTCCGATGATTAATTTGCCACAGTCCTTTAATCTCAGCCAA 1296  
QY  
434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452  
Db  
1297 ACAACAAGCTTTGAA---AAAGACTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1353  
QY  
453 SerMetArgArgGlnThrLeuLeuGlyGly---GlyAspGlyValSerSerThrAsp 471  
Db  
1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413  
QY  
472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491  
Db  
1414 ATTGGTGGTGGCTGGCACTGGCATGGAATGGACTGATATA---GGTGAGGATGGAATA 1470  
QY  
492 LysGluGlyGlyPheLysArgValTyrlleHisGlnGluGlyValProGlySerArgArg 511  
Db  
1471 CAACAAGAGGGGTTTAAAGAGATTTATTTATGAGGAGGAGGTTTGTGCACTCTGCTGT 1530  
QY  
512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531  
Db  
1531 GGATCCATTTGATTCGATTTCCCGGTGAAGC-----GAATTTGTCCAG 1572  
QY  
532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551  
Db  
1573 GCTGTGCTTGGTAAAGCAACCCGCTCTTTACTTCCAAAGGAGCTTATTGATGGACACCCA 1632  
QY  
552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571  
Db  
1633 GTTGGGCTGCAATGGTTCCACCATCTGAGACAGCTTCAAAGGGGCCAAGTTGAAAGCT 1692  
QY  
572 LeuPheGluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGln 591  
Db  
1693 CTTCTTGAACCAAGGGTTAAGCATGTTGTTGGAGTTGGANATACAAATCTTCTCAG 1752  
QY  
592 GlnPheAlaGlyIleAsnGlyValLeuTyrlleThrProGlnIleLeuGluGlnAlaGly 611  
Db  
1753 CAGTTTTTCAAGGATAAATGGGGTCTTATTATACACCTCAATCTTGAAGGCCGCT 1812  
QY  
612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631  
Db  
1813 GTTGAAGTTCTTCTTTCAGATATAGGCATATGGCTCAGATCGGCATCTTCTTATCAGT 1872  
QY  
632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651  
Db  
1873 GCTTTTCAAAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGATGTT 1932  
QY  
652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671  
Db  
1933 TCAGGCAGAGGCGAGTTGCTACTTACTTACAATCCCGTGTGATTGTGTCTACTTATT 1992  
QY  
672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691  
Db  
1993 TTGGTTCATGGAGGCTGTGTAATTTTGGCAATGTGCCCATGCGCAATCTCAACAGTA 2052  
QY  
692 SerValIleValTyrllePheCysPheValMetGlyPheGlyProIleProAsnIleLeu 711  
Db  
2053 TCGGTTGTGGTTTATTTCTGCTGCTTTGTGATGGTTATGACCAATTTCCAAATCTT 2112  
QY  
712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731  
Db  
2113 TGCTCAGAGATTTTCCCACTAGGAGTGGTGGCTCTGTCATTTGCTATCTGTGCAATTAGTG 2172  
QY  
732 PheTrpIleGlyAspIleValThrTyrlleSerLeuProValMetLeuAsnAlaIleGly 751  
Db  
2173 TTCTGGATTGGAGACATCATCATCATCTGCTGTGCTGTGATGCTCGGCTCTTTTAGGA 2232  
QY  
752 LeuAlaGlyValPheSerIleTyrlleAlaValValCysLeuIleSerPheValPhe 771  
Db  
2233 CTTGGTGGTGTATTCGCCATTTACGAGTGTCTTGTGTTTCTCTCGTGGATTTGTGTTT 2292  
QY  
772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791  
Db  
2293 TTGAAGGTTCCAGAAACAAAGGCGATGCCCTTGAAGTCATCTCTGAATCTTCTTCTGTT 2352  
Db







Dd		235	GATAATGCTACCAATCGCCGGGGCTAATGGTTATCATTAAGAAAGAACCTTGCTTGTGGGAACA	294
Qy		96	GluproThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrlieValThr	115
Dd		295	-----ACTATGAAAGGCATTGGTGGGCATGTCCTGATTGGAGCAACGGTAATCACC	348
Qy		116	ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSer	135
Dd		349	ACATGCTCTGGTCTATAGCGGATGGCTCGGTGGCGGCCCATGATGATAATCTCATCT	408
Qy		136	IleLeuTyrrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu	155
Dd		409	GTGCTCTAATTTCTGGGTGGTTTTGGTAGTCTGTGTGTCGCCAAATGTGTATGTGTGTGC	468
Qy		156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Dd		469	TTGGCGAGGCTACTTGATGGATTGGGATTGGCTTGGCTGTGTGACTCTGTGCCGGTCTAT	528
Qy		176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Dd		529	ATATCTCAAACGGCGCGCTCTGAATAAAGGGGTCGTGTAATACGCTTCCTCAGTTCAGT	588
Qy		196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Dd		589	GGCTCTGGAGGAATGTTTTGTGCTACTGATGGTTTTTGGCATGTCATTGAGTCCCGCG	648
Qy		216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Dd		649	CCTAGCTGGAGGCTCATGCTTGGGGTCTGTCTATTCTCTCTGTATTATTGGCATTG	708
Qy		236	ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255
Dd		709	ACCATTTTTTCTTGGCCGAGTCTCTCGGTGGCTGTGTGTCAGCAAAAGAAGATGCTCGAG	768
Qy		256	AlaIlyValValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275
Dd		769	GCTAAGAAGGTGCTCCAAGATTTCGCGGAAGGGAGGATGTGTGAGCGAGATGGCATTG	828
Qy		276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
Dd		829	CTGGTTGAAGTCTCGGATTGGGGTGATACATCTATCGAAGAGTACATAATGGCCCT	888
Qy		296	AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGlyGlnIleThrLeuTyr	315
Dd		889	GCTGACGATGCTGCTGATGTCATGAACATGCAACAGAGAAAGATAAAAATTCGATTATAT	948
Qy		316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335
Dd		949	GSATCCCAACGACGCCCTTCTTGTTATTCAAAACCTGTCTACTGGACAGAGTTCTATTGGC	1008
Qy		336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Dd		1009	-----CTTGGCTCACCATGGGAAGCATCATCAACCAAGCATGCCCTCATG	1056
Qy		356	AspProfileValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
Dd		1057	GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAGTCTCCCGAGACAGAGCAAGA	1116
Qy		374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Dd		1117	GGAAGCATGGGAAGCATCTGTTTTCCAAATTTTTGGAAGCATGTTTCAGCACTGCTGCAGCG	1176
Qy		394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Dd		1177	CATGCTAAAAATTGAACANTGGATGAAGAAGCTTTCAAAAGGNAACGTGAGGACTACATG	1236
Qy		414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Dd		1237	TCAGATCAACCCGTGGGACTCCGATGATAAATTTGACAGTCTCTTTAATCTCACGCGCAA	1296
Qy		434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu---	452
Dd		1297	ACAACAAGCCTTGAA-----AAAGACTTACCTCTCTCTCTCCCATGCGCATCTCTGGC	1353

Qy	453	SerMetArgArgGlnThrLeuLeuGlyGluGly--GlyAspGlyValSerSerThrAsp	471
Db	1354	AGCATGAGCGGTACACATAGTCTCATGCAAGGGTCAGGTGAGCAAGGTCGTATGACAGT	1413
Qy	472	IleGlyGlyGlyTrpGlnLeuAlaTprLysTrpSerGlnLysGluGlyGluAenGlyArg	491
Db	1414	ATTGGTGGTGGCTGGCAACTGGCATGGAAATGACCTGATANA---GGTGAGGATGGNAAA	1470
Qy	492	LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg	511
Db	1471	CAACAAGGAGGGTTTAAAGGATTTATTATACATGAGGAGGAGCTTCTGTCATCTCGTCGT	1530
Qy	512	GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis	531
Db	1531	GGATCATTTGTATCGGATTCGCCGTGAAGGC-----GAAATTTGTCCAG	1572
Qy	532	AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet	551
Db	1573	GCTGCTGCCTTGGTAGCAACACCCGCTCTTTACTCCAAGAGAGCTTATTGATGACACCCCA	1632
Qy	552	SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp	571
Db	1633	GTTGGGCGCTGCANTGGTTCAACCATCTGACAGAGCTTCAAAGGGCCCAAGTTCGAAAGCT	1692
Qy	572	LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln	591
Db	1693	CTTCTTTGAACCGGGTTAAGCATGCTATTGGTTGGAGTTGGAATACAAATACTATTCTAG	1752
Qy	592	GlnPheAlaGlyIleAenGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly	611
Db	1753	CAGTTTTTCAGGGATAAATGGGGTTCTATATACACACCTCAAAATCCCTTGAAGAGGCGCGT	1812
Qy	612	ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer	631
Db	1813	GTTGAAGTTCTTCTTTCAGATATAGGCATTTGGCTCAGAGTCGGCAATCATTTCCCTATCAGT	1872
Qy	632	SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu	651
Db	1873	GCTTTCAACACCTTCTTGATGCTTCCTGTATAGGCGTAGCCATGAAGCTCATGGATGTT	1932
Qy	652	SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle	671
Db	1933	TCAGGCAGAGGAGATTGCTACTTACTACAATCCCGTGTGATTGTGTCACTCATTTATT	1992
Qy	672	LeuValValSerAenLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal	691
Db	1993	TTGGTCATTGGAGACCTGGTAAATTTTGGCAATGTGCCCATGACGCAATCTCAACAGTA	2052
Qy	692	SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAenIleLeu	711
Db	2053	TGGCTTGTGGTTTATTCTGCTGTTGTGATGGTATTATGACCAATTCGCAACATCCTT	2112
Qy	712	CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr	731
Db	2113	TGCTCAGAGATTTTCCCACTAGGCGTGGCGCTCTGCAATGCTATCTGTGCAATTAGTG	2172
Qy	732	PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAenAlaIleGly	751
Db	2173	TTCTGGATTGGAGACATCATCACATATCTCGCTGCCTGTGATGCTCGGCTCTTAGGA	2232
Qy	752	LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe	771
Db	2233	CTTGGTGGTGTATTGCGCATTTACGCAGTTGTTTGTTCATCTCGTGGATATTGTGCTTT	2292
Qy	772	LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaVal	791
Db	2293	TTGAAGGTTCCAGAAACAAAGGCGATGCCCTTTGAAGTCATCTGAATCTTTTCTGTT	2352
Qy	792	GlyAlaLysGlnAlaAlaAla	798
Db	2353	GGAGCAAGCAGCGCTGCTCT	2373



```
Db 1240 GGATGCCAAGTGGCATGGAATAATGCAGGAAGAGAGATGAATCGGGACAGAAAGAA--- 1296
Qy 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
Db 1297 -----GAGGTTTCCAGGATCTCGACGTGGCTCAATT 1329
Qy 515 ValSerLeuProGlyGlyAepValPheGluGlySerGluPheValHisAlaAla 534
Db 1330 GTTTCATTGCTGGTGGTATGATGAACCGGTGAG---GCAGATTTTGTACAAGCGTCTGCT 1386
Qy 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
Db 1387 TTGGTTAGCAACACAGCTCTTTATTCGAAGACCTTCTCAAGAACATACAAATTGGTCCT 1446
Qy 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAepLeuPheGlu 574
Db 1447 GCTATGGTACATCCATCCGNA---ACNACTAAGGGTCAATTTGGCATGATCTTCATGAT 1503
Qy 575 ProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594
Db 1504 CTTGGAGTCAAGCGTGCATTAGTCTAGGAGTTGGACTTCAATACTCTCAGCAGTTCTCA 1563
Qy 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaVal 614
Db 1564 GGCATCAACGGAGTCTTTACTACACACCGCAATCTTGAGCAGCGGGTGTGGGATC 1623
Qy 615 IleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuLeuSerSerLeuThr 634
Db 1624 CTACTATCGACATGGGATAGTTCTTCTCAGCATCTTACTATTAAGTGAATGACA 1683
Qy 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAepLeuSerGlyArg 654
Db 1684 ACCTTTGTGATGTACTCGCAATAGCTGTGGTGAATGAGGCTCATGATCTTCTGTCGA 1743
Qy 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 674
Db 1744 AGGACCTTGTCTTCCACGATACCAATCTGATAGCATCTCTATTTGGTTTAGTAATC 1803
Qy 675 SerAsnLeuIleAepLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
Db 1804 TCNAATCTTGTTCACATGAACACATGTGTCAGCGGTCTTATCAACCGTAACGGTTGTG 1863
Qy 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
Db 1864 CTCTACTTCTGCTTCTTGGTATGGGTTTGGTCTGCTCCAAACATCTCTGTTTCAGAG 1923
Qy 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
Db 1924 ATTTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGCGCATCTCACCTTCTGATC 1983
Qy 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
Db 1984 TGTGACATAATCGTCACTTACAGTCTCCCGTGTCTCTCAATCCATGGACATGAGCTGGT 2043
Qy 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774
Db 2044 GTGTTTGGATGTACGCAATCGTATGTGCAATTTTCATGGGTCTTTGTGTTTCATTAAAGTC 2103
Qy 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
Db 2104 CCGGAACAACTAAGGCATGCCACTTGAAGTCAATCAGAGTCTTCTTCTGTTGGAGCTAGA 2163
Qy 795 GlnAlaAlaAla 798
Db 2164 CAAGCTGAAGCT 2175
RESULT 10
ID ABZ14449
XX ABZ14449 standard; DNA; 2205 BP.
AC ABZ14449;
XX
```

```
DT 21-JAN-2003 (first entry)
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2254.
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
OS Arabidopsis thaliana.
PN WO200216655-A2.
PD 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US026685.
XX 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX (SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
PI Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
DR Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
PT Claim 144; SEQ ID NO 2254; 577pp + Sequence Listing; English.
PS The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
SQ Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,06e-189 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
DB: 6 Gaps: 14

US-10-051-909-32 (1-800) x ABZ14449 (1-2205)
Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 1 ATGAAGGAGCGACTCTCGTTGCTCTCGCGCCACCAATCGGCAATTTCTTACAGGATGG 60
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
Db 61 GACAATGCCACCATTCTGGAGCTATGTTTATATCAACAAAGACTTGAATCTA----- 114
Qy 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 115 ---CCAACTCTGTTCAAGGTCTTGTCTGTTGCTATGTCATGTCGGTGCACGGTCATC 171
Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuLeuLeuSer 134
Db 172 ACGACTTGTCTCAGACCGATATCTGATTTGGTGGCAGACGCCCATGCTCATTTATCA 231
Qy 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154
```

```
Db 232 TCAGTTATGATATTCGTCGCGTTTGATAATGTTGGTCTCCCAATGTCATGTTCTG 291
QY 155 LeuLeuAlaAArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 292 TGGCTTTGCTAGGCTCTTTAATGGGTTTGTCGCGGCTCGCGGTTTACACTTGTCCCTGTT 351
QY 175 TyrIleSerGluIleAlaProSerGluIleAArgGlyLeuLeuAenThrLeuProGlnPhe 194
Db 352 TACATTTCTGAAACCGCTCTCCCGAGATCAGAGACAGTTAAATACTCTCCCTCAGTTT 411
QY 195 SerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 412 CTGGCTCTGGTGGAAATGTTTGTTCATCTGTATGTTTTCATATGTCCTCGCTGAGTGAC 471
QY 215 SerProAspTTPArgIleMetLeuGlyValleuAlaIleProSerLeuPhePheGly 234
Db 472 TCCCTTAGCTGAGAGCCATGCTCGGTGCTCTCGATCCCTCTCTCTTTATTTGTTT 531
QY 235 LeuThrIlePheTyrIleuProGluSerProArgThrProArgThrLeuValSerLysGlyArgMetAla 254
Db 532 CTCACGGTGTATTTATTTGCCCGAGTCTCTCGTTGGCTGGTTAGTAAAGGAAGAATGGAC 591
QY 255 GluAlaLysValLeuGlnLysLeuArgGlyLysAspValSerLysGlyLeuSer 274
Db 592 GAGGCTAAGCGAGTCTTTCAACAGTTATGTGCAGAGAAGATGTTACCGATGAGATGGCT 651
QY 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGly 294
Db 652 TTACTAGTTGAAGGACTAGATAGGAGGAGAAAAACAATGGAAGATCTCTTAGTAAC 711
QY 295 ProAlaThrGluAlaAspAspLeuValThrAspGlyAspLysGlu--GlnIleThr 313
Db 712 TTGGAGGATCATGAAGGTGATGATACATCTGAACCGTTGATGAGGATGACAAATCGG 771
QY 314 LeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db 772 CTTTATGGAACCCACGAGAATCAATCGTACCTTGCTAGACCTGCCAGCAACAA----- 825
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 826 ---AATAGCTACTTGGGCTACGCTCTCCCAACCGAAGCTTAGCAAAACCAAGCATGATC 882
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAenMetProGlnAlaGly 373
Db 883 CTTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAGAATGCCAGAACGAGGC 942
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 943 GGAACACTCGGAGTGGGATTTTCCTCATTTCCGAAAGCATGTTCACTACTCTCCGAT 1002
QY 394 -----HisAlaLysAsnGluLysTrpAsp-----GluGluAenLeuHisArgAsp 408
Db 1003 GCGCCTCAGGTAAACCGGCTCATTGGGAAAGGACATAGAGCCATTACACAAAGAC 1062
QY 409 AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp-----AsnLeu 425
Db 1063 AATGATGATAGTATGCACTGATGTTGTCGCGGTGATGATGATGATGATGATGATGATG 1122
QY 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445
Db 1123 CGTAGCCCCCTTAATGTGTCGCGCACCAACAGCATGAC---AAGGATATGATCCCAAT 1179
QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAsp 465
Db 1180 CCTACAGTGAAGCACCTTTAAGCATGAGACACAGTACGCTTATGCAA---GGCAAC 1236
QY 466 GlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLys 485
Db 1237 GSCGAAAGTAGCATGGGAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1296
QY 486 GluGlyGluAenGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
Db 1297 GAA-----TACAAGAGGATTATCTTAAAGAGATGGA 1329
```

```
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525
Db 1330 GCTGAA---TCTCGCGGTGGCTCGATCATCTATTCCCGAGGTCCGGAT-----GGT 1380
QY 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
Db 1381 GGAGGCAGCTACATTACGCTTCTGCTTGTAAAGCAGATCTGTTCTTGGTCTTAATCA 1440
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
Db 1441 -----GTTCATGGATCCGCCATGGTTCCCGGAGAAAATTTGTCCTCT 1485
QY 566 GlySerArgTTPAspLeuPheGluProGlyValArgAlaLeuLeuValGlyVal 585
Db 1486 GGACACATCTGTGCTGCTCTTCTTGAACCTGGTGTAAAGCTGCTTGGTGTGGTGTGTC 1545
QY 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605
Db 1546 GGCATTTCAATACTCAGCAGTCTTTCAGGTATCAATGGAGTTCTCTACTACACTCTCTCAG 1605
QY 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
Db 1606 ATTCTCGAACGGGCTGGGCTAGATATTCTTCTTCGAGCCTCGGACTAAGTTCATCTCT 1665
QY 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
Db 1666 GCGTCACTCTCATCAGCGGTTTAAACAATTAATCTCATGCTCCCGACCAATTTGCTGGTGC 1725
QY 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 665
Db 1726 ATGAGACTCATGGATGATCCGGAAGAAGTCAATTCTTCTGGACANTCCCGATTCTC 1785
QY 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
Db 1786 ATTGCTCACTTGTGCTCTTGTATCCGGAAGAAGTCAATTCTTCTGGACANTCCCGATTCTC 1845
QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
Db 1846 GCAGCACTCTCCACAGGTTGTGCTGCTCTACTTCTTCTTCTGATGGGTACGGT 1905
QY 706 ProfileProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
Db 1906 CCCATTCCAAACATCTCTGTTCTGAAATCTTCCCAACAAGAGTCCGTGCTCTCTGCATC 1965
QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745
Db 1966 GCATATGCTGCTATGCTCTTTTGGATTGGAGACATTATTGTACGTACTCACTTCCCGTT 2025
QY 746 MetLeuAenAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
Db 2026 CTCTCAGCTCGATCGACTAGTTGGTGTGTTTTCAGCATTTACGCTGCGGTTGGGTATC 2085
QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
Db 2086 TCATGGATCTTCTGTTTACATGAAAGTCCCGAGACTAAAGCATGCTTTTGGAGATTATC 2145
QY 786 ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 2146 ACAGACTACTTTGCTTTGGAGCT---CAAGCTCAAGCTTCTGCT 2187

RESULT 11
ADG87978
ID ADG87978 standard; cDNA; 2205 BP.
XX
AC ADG87978;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP4-upregulated pathogen infection-related gene #420.
XX Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
```



QY	446	GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAsp	465
Db	1180	CCTACAAAGTGAAGCATTAAAGCATGAGACACAGTACGCTTATGCAA--GGCAAC	1236
QY	466	GlyValSerThrAspIleGlyGlyValGlnLeuAlaTrpLysTrpSerGluLys	485
Db	1237	GGCGAAGTAGCATGGGAATTGGTGGTGGCATATGGATATAGATACAAACGAT	1296
QY	486	GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyLeuHisGlnGluGly	505
Db	1297	GA--TACAAAGGTATTATCTTTAAAGAGATGGA	1329
QY	506	ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspValPheGlu	525
Db	1330	GCTGAA--TCTCGCGTGGCTCGATCATCTATTCCCGAGGTCGGAT-----GGT	1380
QY	526	GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly	545
Db	1381	GGAGGCGAGCTACATTCACGCTTCTGCCCTTGTAAAGCAGATCTGTTCTTGGTCTTAAATCA	1440
QY	546	LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys	565
Db	1441	-----GTTTCATGATCGCCATGGTTCCCGGAGAAAATTGCTGCCTCT	1485
QY	566	GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuValGlyVal	585
Db	1486	GGACCACTGTGTCTGCTCTTCTTGAACCTGGTGTAAAGCGTGCCTTGGTGTGGTGC	1545
QY	586	GlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyThrProGln	605
Db	1546	GGCAITTCAAATACTGCAGAGTTTCAGGATTCAAAGGAGTCTCTACTACACCTCTCAG	1605
QY	606	IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer	625
Db	1606	ATTCTCGAACGGGCTGGCTAGATATCTTCTTCGAGCCTCGGACTGAAGTTCCATCTCT	1665
QY	626	AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla	645
Db	1666	CGGTCTATCTCATCAGCGGTTTAAACACATTACTCATGCTCCAGCCATTGCTGGTGC	1725
QY	646	MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeu	665
Db	1726	ATGAGACTCATGATGATCCGGAAGAGTTCATTACTTCTGGAACAATCCAGTCTCTC	1785
QY	666	IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis	685
Db	1786	ATGTCTCACTGTCTGCTCTTGTTCATCAGCGAGCTCATCCACATCAGCAAGTGTGAAC	1845
QY	686	AlaLeuLeuSerThrValSerValIleValTyPheCysCysPheValMetGlyPheGly	705
Db	1846	GCAGCACTCTCCACAGGTTGTGTGGTCTACTCTCTCTCTGATGGGTACGGT	1905
QY	706	ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle	725
Db	1906	CCCATTTCCAAACATCTCTGTTTCCAAATCTTCCCAACAAGAGTCCGGTGTCTGCAATC	1965
QY	726	AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTySerLeuProVal	745
Db	1966	GCCATATGTCTATGCTCTTTTGGATTGGAGACATTATGTCACTACTCACTTCCCGGT	2025
QY	746	MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyAlaValValCysLeuIle	765
Db	2026	CTCTCTCAGCTCGACTAGTGTGTGTGTTCAGCATTTACGCTGCGGTTTGGGTATC	2085
QY	766	SerPheValPheValPheLeuValProGluThrLysGlyMetProLeuGluValIle	785
Db	2086	TCAAGATCTTCTTTCATGAAGTCCCGGAGACTTAAGCATGCTCTTGGAGTTATC	2145
QY	786	ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla	800
Db	2146	ACAGACTACTTTCCTTTCGAGCT---CAAGCTCAAGCTTCTGCT	2187

ABK51966	ABK51966	standard; cDNA; 1692 BP.
AC	ABK51966;	
XX	27-AUG-2002	(first entry)
DE	Soybean cDNA clone ss1.pk0022.f1	encoding sugar transport protein.
XX	Soybean; Arabidopsis thaliana-like	sugar transport protein;
KW	carbohydrate transport; grain filling;	annual field crop; plant;
KW	clone ss1.pk0022.f1; Gene; ss.	
OS	Glycine max.	
XX		
Key	Location/Qualifiers	
FT	9..1469	
CDS	/tag= a	
FT	/partial	
FT	/product= "Portion of a soybean Arabidopsis thaliana-like	
FT	sugar transport protein"	
FT	/note= "This sequence lacks a start codon"	
XX		
PN	US6383776-B1.	
XX	07-MAY-2002.	
PD		
XX	14-APR-1999;	99US-00291922.
PF		
XX	24-APR-1998;	98US-0083044P.
PR		
XX	(DUPO ) DU PONT DE NEMOURS & CO E. I.	
PA		
XX	Allen SM, Hitz WD, Kinney AJ, Tingey SV;	
PI	WPI; 2002-453364/48.	
XX	P-PSDB; AAU97205.	
DR		
XX	New nucleic acid encoding plant sugar-transport proteins, useful for	
PT	preparing transgenic plants with altered carbohydrate distribution.	
XX	Claim 3; Col 41-44; 54pp; English.	
XX		
CC	The present invention relates to the isolation of plant polynucleotide	
CC	sequences encoding an Arabidopsis thaliana-like sugar transport protein	
CC	or Beta vulgaris-like sugar transport protein. The polynucleotide	
CC	sequences are useful for altering the level of sugar transport proteins	
CC	in plants, i.e. for control of carbohydrate transport and distribution in	
CC	plant cells, e.g. during grain filling of annual field crops (e.g. corn,	
CC	rice, soybeans, and wheat), and, for studying carbohydrate flows and	
CC	sugar transport. The polynucleotide sequences can also be used to isolate	
CC	cDNA sequences and genes that encode homologues of the new proteins. The	
CC	present sequence encodes a portion of a soybean Arabidopsis thaliana-like	
CC	sugar transport protein	
XX	Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 U; 0 Other;	
SQ		
Alignment Scores:		
Pred. No.:	2,32e-136	Length: 1692
Score:	1670.50	Matches: 333
Percent Similarity:	79.75%	Conservative: 57
Best Local Similarity:	68.10%	Mismatches: 84
Query Match:	40.99%	Indels: 15
DB:	6	Gaps: 8
US-10-051-909-32	(1-800) x ABK51966	(1-1692)
QY	308	AspLysGluGlnIleThrLeuTyArgProGluGluGlyGlnSerTrpIleAlaArgPro 327
Db	21	GAATAAGATCAAATTAAGTTGATGACCAAGCAAGCCAGTCTCTGGTTCCTAGACT 80
QY	328	SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347

Db	81	GTTCCTGGACCAAAATTCGTGGC-----CTTGATATCTAGGAAGGAAGCATG	128
Qy	348	ValAenGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu	367
Db	129	GCAAAATCCNAGC--AGTCTAGTGGACCCCTAGTGACCCCTCTTTGGTAGTGTACATGAG	185
Qy	368	AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet	387
Db	186	AAGCTCCAGAAAACAGGA-----AGCACCCCTTTTCCACACTTTGGGAGTATG	233
Qy	388	PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg	407
Db	234	TTCAAGTGTGGGGGAAATCAGCCAAAGAAATGAAGATTGGGATGAGGAAGCCCTAGCCAGA	293
Qy	408	AspAspGluGluTrpAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer	427
Db	294	GAGGGTCATGATATATGCTCTCGAT-----GCTGGTGATTCGTGATGACATTTGCAGAGT	347
Qy	428	ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis	447
Db	348	CCATTGATCTCACGTCCTCAAAACACAGAGCTCGAT---AAGGACATACCTCTCATGCCCAT	404
Qy	448	ArgGlySerAlaLeuSerMetArgGlnThrLeuLeu---GlyGluGlyGlyAspGly	466
Db	405	AGTAACTTGCAC---AGCATGAGCAAGGTAGTCTTTTACATGGAAATTCAGGAGAACCC	461
Qy	467	ValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlu	486
Db	462	ACTGGTAGTACTGGGATGGTGGTGGTGGTGGCAGCTAGCATGGAAATGGTCTGAAAGAGAG	521
Qy	487	GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal	506
Db	522	GGCCACAGATGGAAGAAGGAAGGTGGCTTCAAGAGAATATATTTACACCAAGATGGTGGT	581
Qy	507	ProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly	526
Db	582	TCTGGATCTAGACGTGGGTCTGTGGTTTCACTCCCT---GGCGGTGATTTACCAACTGCAC	638
Qy	527	SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu	546
Db	639	AGTGAGGTTGTACAGCTGCTGCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	698
Qy	547	AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly	566
Db	699	ATGGCTCAACGGCCAGTTGGACAGCTATGATTCCCTCTGAAACAAATTCGCAAAAGG	758
Qy	567	SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly	586
Db	759	CCAAGTTGGAGTGAATCTTTTGAACCTGGGGTGAAGCATGCATTTGATTTGGGGGTGGGA	818
Qy	587	IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle	606
Db	819	ATGCAAATCTTTCAGCAGTCTCTGGGTATAAATGGGGTCTCTACTATAGCCCTCAAAAT	878
Qy	607	LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla	626
Db	879	CTTGAGCAGCAGGTGTGGTTATCTTCTTTCAGCCCTAGGCCCTGGTCTTACTTCTTCA	938
Qy	627	SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet	646
Db	939	TCCCTTCTTATAGTGGGGTGACAACTTGTGTGATGCTTCTGTATAGCCATTTGCCCATG	998
Qy	647	LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIle	666
Db	999	AGGCTCATAGGATATTTTCAGGCAGAAAGACTTTGTGCTCAGTACAAATCCCGCTCTAATA	1058
Qy	667	AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla	686
Db	1059	GCACCTCTTCTCATATTAGTCTGGGAAGTCTGTGGATTTGGGATTCCTGCAAAATGCA	1118
Qy	687	LeuLeuSerThrValSerValIleValTyrPheCysPheValMetGlyPheGlyPro	706

Db	1119	TCANCTCAACCACTAGTGTTATTGCTCTATTCTTCTGTTCTTTCTTGTCTATGGGATTTGGACCA	1170
Qy	707	IleProAsnIleLeuCyAsAlaGluIlePheProThrArgValaArgGlyLeuCysAlaA	726
Db	1179	ATTCTTAATATACCTTTGTGCAGAGATCTTCCCACTCGAGTTCTGTGCTCTCTGCATTGCT	1238
Qy	727	IleCysAlaPheThrPheThrIleGlyAspIleValThrTyrSerLeuProValMet	746
Db	1239	ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTCACTACACACTCCCAAGTTATG	1298
Qy	747	LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer	766
Db	1299	CTCAATCTCTAGGCCCTCGCTGGTGTTTTTGGTATTTATGCTGTGTTGCTTCATAGCA	1358
Qy	767	PheValPheValPheLeuLeuValProGluThrIleGlyMetProLeuGluValIleThr	786
Db	1359	TGGGTGTTTGTCTTTTGAAGTTCCAGAAACCAAGGCGATGCCACTGGAAGTGATCAT	1418
Qy	787	GluPhePheAlaValGlyAlaLysGln	795
Db	1419	GAGTTCCTCTCTGTCGGAGCAAAACAG	1445
RESULT	13		
ABX93202			
ID	ABX93202	standard; cDNA; 1692 BP.	
XX	ABX93202;		
XX	29-MAY-2003	(first entry)	
XX	cDNA encoding soybean sugar transport protein #2.		
XX	Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;		
KW	plant sugar transport protein; carbohydrate transport; soybean;		
KW	carbohydrate distribution; plant; gene; ss.		
XX	Glycine max.		
XX	US2002178468-A1.		
XX	28-NOV-2002.		
XX	17-JAN-2002; 2002US-00051902.		
XX	24-APR-1998; 98US-0083044P.		
PR	14-APR-1999; 99US-00291922.		
XX	(ALLE/) ALLEN S M.		
PA	(HITZ/) HITZ W D.		
PA	(KINN/) KINNEY A J.		
PA	(TING/) TINGEY S V.		
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;		
XX	WPI; 2003-340957/32.		
DR	P-PSDB; ABU08330.		
XX	Novel plant sugar transport proteins and nucleic acid encoding the		
PT	protein useful for producing transgenic plants having altered levels of		
PT	sugar transport protein.		
XX	Claim 2; Page 22-23; 56pp; English.		
XX	The present invention relates to the isolation of Arabidopsis thaliana-		
CC	like or Beta vulgaris-like sugar transport proteins, and the		
CC	polynucleotide sequences encoding them. The plant sugar transport		
CC	proteins of the invention have been isolated from corn, rice, soybean,		
CC	and wheat. The polypeptides of the invention may be used for altering the		
CC	level of expression of a sugar transport protein in a host cell, by		
CC	transforming a host cell with a chimeric construct encoding all, or a		
CC	portion of the sugar transport protein, in sense or antisense		
CC	orientation. Particularly, the polypeptides may provide a means to		
CC	control carbohydrate transport and distribution in plants. ABX93198-		







Qy	462	GlulGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys	481
Db	12	GAAGGGTGGGGAGCGAGTCAGCAGCACTGGTATTGGTGGGGGGTGGCAACTCGCATGGAA	71
Qy	482	TrpSerGluLysGluGlyGluAlaGngLysGlyGlyGlyGlyGlyValThrLeu	501
Db	72	TGGTCGGAGCGACAAGCGCAGGATGGCAAGGAAGGAGGCTTCAAAAGAATCTACTTG	131
Qy	502	HisGlnGluGlyValProGlySerArgGlySerIleValSerLeuProGlyGlyGly	521
Db	132	CACCAAGAGGGGGTGGCCGCACTCAAGAAGGGGGCTCTGTGTTCCTCTCTGGTGGGGGT	191
Qy	522	AspValPheGlu--GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla	540
Db	192	GAATCCACCAAGGGGGTACTGGTTTATACAGTCGTCTGTTGGTAAGCCACTCGGCT	251

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:36:34 ; Search time 322.938 Seconds  
(without alignments)  
4054.730 Million cell updates/sec

Title: US-10-051-909-32  
Perfect score: 4075  
Sequence: 1 IRSGSLAVQTFFTPDLRR.....PLEVTEFFAVGAKQAANKA 800

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlp  
-Q=/cgn2\_1/USPTO pool p/US10051909/runat 13042005 074038 14080/app query.fasta\_1.1678  
-DB=Issued\_Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0 1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN 1 141 @runat 13042005 074038 14080 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	2824	3	US-09-291-922-1
2	2674	65.6	2601	3	US-09-291-922-7
3	1670.5	41.0	1692	3	US-09-291-922-9
4	1424	34.9	1487	3	US-09-291-922-13
5	943	23.1	1009	3	US-09-291-922-15
6	562	13.8	870	3	US-09-291-922-5
7	496	12.2	1853	3	US-09-291-922-23
8	494.5	12.1	1914	3	US-09-291-922-21
9	491.5	12.1	1914	3	US-09-291-922-19
10	489	12.0	1872	3	US-09-291-922-27
11	463.5	11.4	2089	3	US-09-291-922-25
12	447	11.0	1431	4	US-09-489-039A-4762

13	440.5	10.8	1545	4	US-09-489-039A-4731	Sequence 4731, Ap
14	427	10.5	1506	4	US-09-489-039A-4560	Sequence 4560, Ap
15	426	10.5	1566	4	US-09-489-039A-2378	Sequence 2378, Ap
16	418.5	10.3	1668	4	US-09-614-221A-420	Sequence 420, Appl
17	412	10.1	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
18	368	9.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
19	355	8.7	2856	3	US-09-643-597-135	Sequence 135, Appl
20	355	8.7	2856	4	US-09-480-884A-135	Sequence 135, Appl
21	355	8.7	2856	4	US-09-542-615A-135	Sequence 135, Appl
22	355	8.7	2856	4	US-09-606-421B-135	Sequence 135, Appl
23	355	8.7	2856	4	US-09-221-107-135	Sequence 135, Appl
24	355	8.7	2856	4	US-09-466-396A-135	Sequence 135, Appl
25	355	8.7	2856	4	US-09-476-496A-135	Sequence 135, Appl
26	355	8.7	2856	4	US-09-630-940B-135	Sequence 135, Appl
27	355	8.7	2856	4	US-09-285-479-135	Sequence 135, Appl
28	354.5	8.7	3000	2	US-08-928-692-9	Sequence 9, Appl1
29	354.5	8.7	3000	3	US-09-339-972-9	Sequence 9, Appl1
30	352.5	8.7	1626	4	US-09-614-221A-521	Sequence 521, Appl
31	346	8.5	5228	4	US-09-919-039-216	Sequence 216, Appl
32	343	8.4	1776	4	US-09-679-686B-11	Sequence 11, Appl
33	338.5	8.3	1943	4	US-09-774-528-168	Sequence 168, Appl
34	334.5	8.2	5227	4	US-09-919-172-79	Sequence 79, Appl
35	332.5	8.2	2592	3	US-09-591-025-8	Sequence 8, Appl1
36	328.5	8.1	2592	4	US-09-894-927B-8	Sequence 8, Appl1
37	322	7.9	1859	4	US-09-248-796A-6610	Sequence 6610, Ap
38	313.5	7.7	1704	4	US-09-614-221A-534	Sequence 534, Appl
39	312	7.7	1675	4	US-09-679-686B-1	Sequence 1, Appl1
40	312	7.7	1695	4	US-09-614-221A-443	Sequence 443, Appl
41	307.5	7.5	1704	4	US-09-614-221A-96	Sequence 96, Appl
42	302	7.4	2298	4	US-09-248-796A-5965	Sequence 5965, Ap
43	301.5	7.4	3915	4	US-09-023-655-1104	Sequence 1104, Ap
44	293	7.2	987	4	US-09-710-279-1941	Sequence 1941, Ap
45	293	7.2	3065	4	US-09-710-279-3697	Sequence 3697, Ap

ALIGNMENTS

RESULT 1  
US-09-291-922-1  
; Sequence 1, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 2824  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (29)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (622)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (636)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (638)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (669)

FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (771)  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (822)  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (856)  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (889)  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (896)  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (944)  
 US-09-291-922-1

Alignment Scores:  
 Prod. No.: 0 Length: 2824  
 Score: 3517.00 Matches: 684  
 Percent Similarity: 95.18% Conservative: 27  
 Best Local Similarity: 91.57% Mismatches: 34  
 Query Match: 86.31% Indels: 2  
 DB: 3 Gaps: 2

US-10-051-909-32 (1-800) x US-09-291-922-1 (1-2824)

QY 56 MetSerGlyAlaValLeuValAlaLeValAlaSerIleGlyAsnLeuGlnGlyTrp 75  
 Db 238 ATGGGGGGCGCGGTGATGTCGCCATCGCGGCTCTATCGGCACTTCTCGAGGGCTGG 297  
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTrpIleLysLysGluPheGlnLeuGlnAsn 95  
 Db 298 GACAAATGCGACAATTCCTGGAGCGCTCTGTATACATAAAGAAAGAAATTCACATCGCAGGC 357  
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
 Db 358 GAGCCTCTGATCGAAGGCTCATCGTCGCCATGTTCTCATTTGGGGCAACAGTCATCACA 417  
 QY 116 ThrPheSerGlyProLeuSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
 Db 418 ACATCTCGGGGCGCAAGGCTGACTGCGTGTGGTAGGAGGCCCATGCTGCTCGCCTCGGCT 477  
 QY 136 IleLeuThrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValThrValLeuLeu 155  
 Db 478 GTCTCTACTTCTGCTAGTGGCTGGTGATGCTTTGGGGCCCAATTGTGTACATCTTGCTC 537  
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
 Db 538 CTCGCAAGGCTCATGATGGTTTCGATATCGGTTTGGCGGTGCACACTTGTTCCTCTCTAC 597  
 QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194  
 Db 598 ATCTCCGAAACTGCACCGCACGAGANATCTTGGGGCTGNTNGAACACAGTTCGCCGAGTTC 657  
 QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213  
 Db 658 ATTGGGGTCAGAGGAGGATGTTCTCTCTCTACTGATGCTGTTGGGATGTCCTCATG 717  
 QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233  
 Db 718 CCNAAACCTGATTGGAGGCTAGCTTGGAGTCTCTGATCCCTCATCTTATNACTTT 777  
 QY 234 GlyLeuThrIlePheTyrLeuProGluSerProAspTrpLeuValSerIleGlyArgMet 253  
 Db 778 GGACTGACTGCTCTACTTGTCTGCTGATCACCAGGTGCTGTWAGCNAAGAGAGATG 837  
 QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273  
 Db 838 GCGAGGCGAAGAGTGNTGCAAGGCTGCGGGGAAGAGAGATGCTCTCANGGGGAGANG 897

QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293  
 Db 898 GCTCTTCTAGTTGAAGGTTTGGGGTCGGTAAAGATACACGATATTNNAGAGTACATCAT 957  
 QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313  
 Db 958 GGACCTGCCACCGAGGCGCGATGATCTTGTAACTGACGGTGATTAAGGAACAATAACA 1017  
 QY 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333  
 Db 1018 CTTTATGGCCTGAAGAAGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG 1077  
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353  
 Db 1078 CTTGGAAGTGTGCTTCTTCTGCACTCTGTCATGGAGCATGGTGAACAGAGTGTACCC 1137  
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGlyAsnMetProGlnAlaGly 373  
 Db 1138 CTTATGGATCCGATTTGTGACACTTTTGGTAGTGTCATGAGAAATATGCCCTCAAGTGA 1197  
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
 Db 1198 GGAAGTATGAGGAGCACATTTGTTCCAAACTTTTGGAAATGTTTCAGTGTACAGATCAG 1257  
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413  
 Db 1258 CATGCCAAAATGAGCAGTGGGATGAAGAGAATCTTCTATAGGGATGACGAGAGTACGCA 1317  
 QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433  
 Db 1318 TCTGATGGTCGAGGAGGTGACTATCAGGACAATCTCCATAGCCCATTTCTGTCCAGGCAG 1377  
 QY 434 AlaThrGlyAlaGluGlyAspIleValHisGlyHisArgGlySerAlaLeuSer 453  
 Db 1378 GCACAGGTGCGAAGGGAGAGCAATTTGTCACCATGTTCCCGTGGAGTCTTTGAGC 1437  
 QY 454 MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyAspGlyValSerSerThrAspIleGly 473  
 Db 1438 ATGAGAAGCAAAAGCCTCTTATGGGAGGTGGAGATGTTGTGAGCGACCTGATATCGGT 1497  
 QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlyGluAsnGlyArgLysGlu 493  
 Db 1498 GGGGGATGCGACGCTTCTTGGAAATGGTCAGAGAAGGAGTGAAGATGGTAGAAAGAA 1557  
 QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513  
 Db 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTGGCTCAAGAGGGGCTCA 1617  
 QY 514 IleValSerLeuProGlyGlyGlyAspValPheGlyGlySerGluPheValHisAlaAla 533  
 Db 1618 ATTGTTTCACTTCCCGTGGTGGCATGTTCTTGAAGGTTAGTGTGATGTTGTATGCTGCT 1677  
 QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
 Db 1678 GCCTTAGTAAGTCAGTCAGACACTTTTCTCAAGGGTCTTGTCTGAACCCAGCATGTGAT 1737  
 QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573  
 Db 1738 GCTGCCATGTTTCAACCATCTGAGGTAGCTGCAAGAGTTCACGTTGGAAAGATTTGTT 1797  
 QY 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593  
 Db 1798 GAACCTGGAGTAGGCGTCCCTGTTAGTGGTGTGGAAATTCAGATCTTCAACAGTTT 1857  
 QY 594 AlaGlyIleAsnGlyValLeuTyrTrpProGlnIleLeuGluGlnAlaGlyValAla 613  
 Db 1858 GCTGGAATAAAGCGTGTCTGTACTATATACCCCAAAATTTCTTGAGCAAGCTGTGTGGCA 1917  
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuLeuSerLeu 633  
 Db 1918 GTTATTTCTTCCAAATTTGGTCTCTCAGCTCGCATCAGCATCCATCTTGTATCATGTTCTCTC 1977

QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653  
 Db 1978 ACTACCTTACTAATAGTCTCTGTCATGGCTTTGCCATGCTGCTTATGATCTTTCCGGA 2037  
 QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
 Db 2038 AGNAGGTTTTGCTGCTAGGCACAAATCCAAATCTTGATGATCTCTAGTTATCTCGTT 2097  
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
 Db 2098 GTGTCCAACTAATGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157  
 QY 694 IleValThrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713  
 Db 2158 ATCGTCTACTTCGCTCGCTGCTGATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217  
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733  
 Db 2218 GAGATCTTTCCAAACAGGGTTCGTGGCTCTGTATTGCCATTTGTGCTTTACATTTCTGG 2277  
 QY 734 IleGlyAspIleIleValThrTySerLeuProValMetLeuAsnAlaIleGlyIleAla 753  
 Db 2278 ATCGGAGATATCATGCTCACCTACAGCTTCCTGTGATGCTGAATGCTATTGGACTGGCG 2337  
 QY 754 GlyValPheSerIleTyAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
 Db 2338 GGTGTTTTTCAGCATATATGACGTCGTATGCTGATTTCCCTTTGTGTGCTTCCTTAAG 2397  
 QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793  
 Db 2398 GTCCCTGACACAAAGGGATGCCCTTGAGGTATTACCGAATTCCTTGCGATTGTTGGCG 2457  
 QY 794 LysGlnAlaAlaLysAla 800  
 Db 2458 AAGCAGCGGCTCCAAAGCC 2478  
 RESULT 2  
 US-09-291-922-7  
 ; Sequence 7, Application US/09291922  
 ; Patent No. 6383776  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Steve  
 ; APPLICANT: Hitz, Bill  
 ; APPLICANT: Kinney, Tony  
 ; APPLICANT: Tingey, Scott  
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins  
 ; FILE REFERENCE: BB-1163  
 ; CURRENT APPLICATION NUMBER: US/09/291,922  
 ; CURRENT FILING DATE: 1999-04-14  
 ; EARLIER APPLICATION NUMBER: 60/083,044  
 ; EARLIER FILING DATE: April 24, 1998  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 7  
 ; LENGTH: 2601  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 US-09-291-922-7  
 Alignment Scores:  
 Pred. No.: 7,77e-273 Length: 2601  
 Score: 2674.00 Matches: 522  
 Percent Similarity: 82.20% Conservative: 92  
 Best Local Similarity: 69.88% Mismatches: 115  
 Query Match: 65.62% Indels: 18  
 DB: 3 Gaps: 8  
 US-10-051-909-32 (1-800) x US-09-291-922-7 (1-2601)  
 QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
 Db 175 ATGAAGAGGTGCCGCTCTGTTGCTATTTCGCGCTTCCATTGGTAAATTTCTCTCCAAGGATGG 234

QY 76 AspAsnAlaThrIleAlaAlaValLeuTyIleLysValGluPheGlnLeuGlnAsn 95  
 Db 235 GATAATGCTTACCATCGCCGGGCTAATGTTTACATTAAAGAAAGACCTTCTCTTTGGGAACA 294  
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
 Db 295 -----ACTATGAAAGGCTTGTGGTGGGCATGTCCTGATTGGAGCAACGGTAATCACCC 348  
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
 Db 349 ACATGCTCTGTGCTTATAGCGATTGGCTCGGTGGCGAGCCCATGATGATGATTAATCTCATCT 408  
 QY 136 IleLeuTyThrPhePheSerClyLeuIleMetLeuTrpSerProAsnValTyValLeuLeu 155  
 Db 409 GTGCTCTATTTCTTGGGTGTTTGGTGATGCTGTGTGTCCTGCTGCTCTTGTCCCGGCTCAT 528  
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTy 175  
 Db 469 TTGGCGAGGCTACTTGTGATGATTGGGATTGGCTTGTGCTGCTCTTGTCCCGGCTCAT 528  
 QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
 Db 529 ATATCTGAAACGGCGCGCTCGAAATTAAGGGGTGCTTGAATACGCTTCTCTCAGTTTCAGT 588  
 QY 196 GlySerGlyClyMetPheLeuSerTyCysMetValPheGlyMetSerLeuSerProSer 215  
 Db 589 GGCTCTGGAGGAATGTTTGTGCTACTGTATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648  
 QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
 Db 649 CTAGCTGGAGGCTCATGCTTGGGTTCTGCTATTCCTTCTCTCTCTCTCTCTCTCTCTCTCT 708  
 QY 236 ThrIlePheTyLeuProGluSerProArgTrpLeuValSerIleGlyArgMetAlaGlu 255  
 Db 709 ACCATTTTTTCTTGCCGAGTCTCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 768  
 QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerClyGluLeuSerLeu 275  
 Db 769 GCTAAGAAGTGTCTCCAAAGATTGCGCGAAGAGGATGTGTGAGCGGAGATGGGATGG 828  
 QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTrpIleIleGlyPro 295  
 Db 829 CTGGTTGAAGGTCTCGGATTTGGGTTGATACATCTATCGAAGAGTACATATTTGGCCCT 888  
 QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTy 315  
 Db 889 GCTGACGATGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948  
 QY 316 GlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
 Db 949 GGATCCCAAGCAGGCGCTTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTTCTATTGGC 1008  
 QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
 Db 1009 -----CTTGGCTCACCATGAGAGCATCATCAACCAAGCATGCCCTCATG 1056  
 QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373  
 Db 1057 GATCCTCTGTCACACTGTTTGGTAGATTATGAGAAGCTCCCGACAGAGCAAGA 1116  
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
 Db 1117 GGAAGCATGCGAAGCAGCTCTGTTTCCAAATTTTGGAAGCATGTTTCAGCACTGCTGAGCCG 1176  
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGluTyAla 413  
 Db 1177 CATGCTAAATTTGAACATGGGATGAAGAAAGCTTACAAAGGAAGCTGAGGACTACATG 1236  
 QY 414 SerAspGlyAlaGlyGlyAspTyTrpGluAspAsnLeuHisSerProLeuLeuSerArgGln 433  
 Db 1237 TCAGATGCAACCCGTTGGGACTCCGATGATTAATTTGCACAGTCTCTTTAATCTCACGCCAA 1296  
 QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452

```

Db      1297 ACAACAAGCCTTGAA---AAAGACTTACCTCTCTCTCTCCCATGGCAGTATCCTTGGC 1353
Qy      SerMetArgGlnThrLeuLeuGlyGly---GlyAspGlyValSerSerThrAsp 471
Db      1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413
Qy      IleGlyGlyGlyThrLeuLeuAlaThrPheGluLeuGlyGlyGluLeuAsnGlyArg 491
Db      1414 ATGTGTGGTGGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1470
Qy      LysGluGlyGlyPheGlyArgValThrLeuHisGlnGluGlyValProGlySerArgArg 511
Db      1471 CAACAAGGAGGTTTAAAGAGATTTATTTACATGAGGAGGAGTTCTGATCTGTCGCT 1530
Qy      GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
Db      1531 GGATCCATGTATGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTC 1572
Qy      532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db      1573 GCTGTCGCTTGGTAAGCAACCCGCTCTTTACTTCAAGGAGCTTATTTGATGGACACCCA 1632
Qy      552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
Db      1633 GTTGGCGCTGCATGTTTCAACCATCTGAGCAGCTTCANAGGGCCCAAGTTGGAAAGCT 1692
Qy      572 LeuPheGluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGln 591
Db      1693 CTTCTTGAACCAAGGGTTAAGCATGATGTTGTTGAGTTGGAATACAAATATCTTACAG 1752
Qy      592 GlnPheAlaGlyIleAsnGlyValLeuThrProGlnIleLeuGluGlnAlaGly 611
Db      1753 CAGTTTTCAGGATAAATGGGGTCTATATTACACACCTCAATTCCTTGAAGAGCGCGT 1812
Qy      612 ValAlaValIleLeuSerLysPheGlyLeuSerBerAlaSerAlaSerIleLeuIleSer 631
Db      1813 GTTGAAGTCTTCTTTCAGATATAGGCATTCGGCTCAGAGTCGGGATCATCTTATCAGT 1872
Qy      632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db      1873 GCTTTTCACAACCTTCTGATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGGATGT 1932
Qy      652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db      1933 TCAGGCAAGAGCAGTTCCTACTTACTACATCCCGTGTGATTTGCTCATCTCATTTATT 1992
Qy      672 LeuValValSerAsnIleLeuLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db      1993 TTGGTCATTTGAGAGCTGTGTAATTTTGGCAATGTCGCCCATGCAATCTCAACAGTA 2052
Qy      692 SerValIleValThrPheCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db      2053 TCGGTGTGTGTATTTCTGCTGCTTTGTGATGGGTTATGACCAATTCCAACATCCTT 2112
Qy      712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db      2113 TGCCTCAGAGATTTTCCCACTAGGTTGGCTGGCTCTGCTATGCTATCTGTGCATTAGT 2172
Qy      732 PheTrpIleGlyAspIleIleValThrThrSerLeuProValMetLeuAsnAlaIleGly 751
Db      2173 TTCTGGATTGGAGACATCATCATCATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2232
Qy      752 LeuAlaGlyValPheSerIleThrAlaValValCysLeuIleSerPheValPhePhe 771
Db      2233 CTTGTGTGTGTATTTGCCATTTACGAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2292
Qy      772 LeuLysValProGluThrLysGlyMetProGluGluValIleThrGluPhePheAlaVal 791
Db      2293 TTGAAGGTTTCCAGAACCAAGGCGATGCCCTTTGAATCACTCTCTGAATTTCTTTCTGTT 2352
Qy      792 GlyAlaLysGlnAlaAlaAla 798

```

```

Db      2353 GGAGCAAAAGCAGGCTGCTTCT 2373
RESULT 3
US-09-291-922-9
; Sequence 9, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Kingsy, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-9

Alignment Scores:
Pred. No.: 1,11e-166 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 3 Gaps: 8

US-10-051-909-32 (1-800) x US-09-291-922-9 (1-1692)
Qy      308 AspLysGluGlnIleThrLeuThrGlyProGluGlyGlnSerTrpIleAlaArgPro 327
Db      21 GAAAGAAGATCAAAATTAAGTTGTATGACCAAGCAAGCCAGTCTCTGGTTGCTAGACCT 80
Qy      328 SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347
Db      81 GTTGTGGACCAAAATCTGTTGGC-----CTTGATCTAGGAAAGGAAGCATG 128
Qy      348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367
Db      129 GCAAAATCCAAGC---AGTCTAGTGGACCTCTAGTACCCCTCTTTGGTAGTGTACATGAG 185
Qy      368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db      186 AAGTCCCAAGCAACAGGA-----AGCACCTTTTCCACACTTTTGGGAGTATG 233
Qy      388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407
Db      234 TTCAAGTTTGGGGGAAATCAGCAAGGAATGAGATTGGGATGAGGAAGCTAGCCAGA 293
Qy      408 AspAspGluGluThrAlaSerAspGlyAlaGlyAspThrGluAspAsnLeuHisSer 427
Db      294 GAGGTTGATGATATGTTCTCTGAT-----GCTGTGATTTCTGATGACAAATTTGCAGAGT 347
Qy      428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHis 447
Db      348 CCATTTGATCTCAGTCAAAACACGAGTCTGGAT---AAGGACATACCTCTCTCATGCCCAT 404
Qy      448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly 466
Db      405 AGTAACCTTGCA---AGCATGAGGCAAGGTAGTCTTTTACATGGAAATTCAGGAGAACCC 461
Qy      467 ValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlu 486
Db      462 ACTGGTAGTACTGGGATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 521
Qy      487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValThrLeuHisGlnGlyVal 506

```

```

Db 522 GGCCAGATGGAAGAGGAGTGGCTCAAGAGATATATTTACACCAAGATGGTGGT 581
Qy 507 ProGlySerArgGlySerIleValSerLeuProGlyGlyAspValPheGluGly 526
Db 582 TCTGGATCTAGACGTGGTCTGGTCTTCACTCCCT---GGCGGTGATTATACCACTGAC 638
Qy 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
Db 639 AGTGAGGTGTACAGGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 698
Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566
Db 699 ATCGCTCAACGGCCAGTGGACAGCTATGATTCATCCCTCTGAAACAATTGCAAAAGG 758
Qy 567 SerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuValGlyValGly 586
Db 759 CCAAGTTGGAGTGCATCTTTTGAACCTGGGTGAAGCATGCAATTGATTGGGGGTGGGA 818
Qy 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606
Db 819 ATGCAAAATCTTCAGCAGTTCTCTGGTATAAATGGGGTCCCTCTACTATAGCCCTCAAAT 878
Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerAlaSerAla 626
Db 879 CTTGACAGCAGCGTGTGGTATCTCTTTCAAGCTAGCCCTTGGTCTTACTTCTTCA 938
Qy 627 SerIleLeuIleSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCCTTTCTTATTAGTCGGGTGACAACCTTGTGATGCTTCTTGTATAGCCATTGCCATG 998
Qy 647 LeuLeuMetAspLeuSerGlyArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db 999 AGGCTCATGGATATTCAGGCAGAGGACTTTGCTGCTCAGTACAAATCCCGCTCTAATA 1058
Qy 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTCTCATATATTAGTCTGGGAAGTCTTGGATTGGATCCACTGCAAAATGCA 1118
Qy 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATAGTATTATGCTATTCTTCTGTTCTTCTGATGGGATTTGGACCA 1178
Qy 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
Db 1179 ATTCTTAATATACTTTGTGCAGAGATCTTCCCACTCGAGTCTGCTGCTCTGCTGCT 1238
Qy 727 IleCysAlaPheThrPheTrrIleGlyAspIleValThrTyrSerLeuProValMet 746
Db 1239 ATTGTGCCCTTACCCTTTGGATCTGTGATATCATTTGTACCTACACCTCCCGATTATG 1298
Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766
Db 1299 CTCAAATCTGAGCCCTCGTGGTGTGTTTGGTATTATGCTGCTGCTGCTTCATAGCA 1358
Qy 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTGTTCTCTTTTGAAGTTCAGAAACCAAGGGCATGCCACTGGGAAGTGATCAT 1418
Qy 787 GluPhePheAlaValGlyAlaLysGln 795
Db 1419 GAGTTCTTCTCTGTGGAGCAAAACAG 1445

```

# RESULT 4

```

US-09-291-922-13
; Sequence 13, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
; US-09-291-922-13
Alignment Scores:
Pred. No.: 1,27e-140 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 3 Gaps: 2
US-10-051-909-32 (1-800) x US-09-291-922-13 (1-1487)
Qy 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys 481
Db 12 GAGGGTGGGAGGAGCAGTCAAGCAGCTGGTATTGGTGGGGGTGGCAACTCGCATGGAAA 71
Qy 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501
Db 72 TGGTCGGAGCGACAAGCGCAGGATGGCAAGAGAGAGGAGGCTTCAAAAGAAATCTACTTG 131
Qy 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAGAGGGGGGTGGCCGACTCAAGAGGGGCTCTGTTGTTTTCACCTTCTCTGGTGGGGT 191
Qy 522 AspValPheGlu---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGCGAGTGGTATTATACATGCTGCTGCTTGGTAGCCACTCGGCT 251
Qy 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGGATCTTATGGAAGAGCGTATGGCGCGCGTCCAGCCATGATTATCCA 311
Qy 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGGAGGCGAGCTCCCAAGAGTTCAATCTGGAAGAGATCTGTTTGAAGCTGTGTGAGCGT 371
Qy 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTTGTCGGTGTGGATTCAAGATGCTTCAGATGCTTCAGCAGTTTCTGGAATAATGGAGTT 431
Qy 600 LeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCTACTACTCTCTCAAAATTTCTGGAGCAAGCTGGTGTGGCTGTTCTTCTTCCAAATCTT 491
Qy 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCCTCAGTTACAGCATCCATCTTGATCAGTTCTCTCACCACCTTACTCATGCTC 551
Qy 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeu 659
Db 552 CCAAGCATTTGGGTAGCCATGACATTCATGATATATCTGGAAGAAGGTTTCTTGCTACTG 611
Qy 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAop 679
Db 612 GGCACAATTTCCCATCTTTGATAGCATCCCTAAATGTTTGGGTGGTGGTCAATGTTATCAAC 671
Qy 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGGTGGCCCGCTGCTCTCCACAGTTAGCGTCAATGTTACTTCTTCTGCTGC 731
Qy 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTATGGGCTTTGGCCCGATCCCAACATTTCTATGTGCAGAGATTTTCCCACACAGA 791

```





```
QY 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheVal 768
|||
DB 248 GCATTTGGACTCGCTGGAGTGTTCGAATCTACGCAGTGTCTGCATATCTGGCTTTCCTG 307
|||
QY 769 PheValPheLeuValProGluThrLysGlyMetProLeuGluValIleThrGluPhe 788
|||
DB 308 TTGTCTTCATGAAGTGGCGGAGACAAAGGGCATCGCTTTGAAGTCATCACCGAGTTC 367
|||
QY 789 PheAlaValGlyAlaLysGlnAla 796
|||
DB 368 TTCCTGTCGGAGCAAAAGCAGGCC 391
|||
RESULT 7
US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23
Alignment Scores:
Pred. No.: 5,6e-42 Length: 1853
Score: 496.00 Matches: 165
Percent Similarity: 36.39% Conservative: 109
Best Local Similarity: 21.91% Mismatches: 189
Query Match: 12.17% Indels: 290
DB: 3 Gaps: 18
US-10-051-909-32 (1-800) x US-09-291-922-23 (1-1853)
QY 59 AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAla 78
|||
DB 257 GCTTTTGTGCTGTATGCTGGCTCCATGACTTCCATCTTGTCTTGTATGATATGGA 316
|||
QY 79 ThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGlu----- 96
|||
DB 317 GTGATAGTGGACGACCCATATACATAAAGGAGCCTGAAAGTCTCGGACGAGCAATC 376
|||
QY 97 ProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 116
|||
DB 377 GAGATCTGCTCGGAATCATCAACCTATATCTCTGATAGCTCATGTCTC----- 427
|||
QY 117 PheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSerIle 136
|||
DB 428 ---GCCGGCAGAACCTCCGACTCGATAGGTCCTCCCTACACGATTTTCGCCGGCACC 484
|||
QY 137 LeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 156
|||
DB 485 ATCTTCTTTCGGAGCACTTCTCATGGGTTCCTCCCCANTTATTCCTTCTCATGTTT 544
|||
QY 157 AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIle 176
|||
DB 545 GGCGGTTTCGCTGGCATTCGATCGGCTAGCCCTCATGATAGCCCGCTACACC 604
|||
QY 177 SerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGly 196
|||
DB 605 GCCGAGGTCTCCCGGCTCTCTCGTGGCTTCTCACTTCCTTCCTTGAGGTATTATT 664
|||
```

```
QY 197 SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSer---LeuSerProSer 215
|||
DB 665 AATGGAGGGATATTAAATTGGATACATATCAAACTATATGCTTTTCGAAGCTGCACCTAAAG 724
|||
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
|||
DB 725 GTGGGATGGCAATGATGCTTGGAGTTGTGTGATACCTTCGGTA---CTCCTTAACAGTA 781
|||
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
|||
DB 782 GGAGTGTTCGGCATGCCGAGTCCCCCAAGTGGCTTGTGATGAGGGTTCGTTGGGAGAG 841
|||
QY 256 AlaLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275
|||
DB 842 GCAAGAAAAAGTGTCTTAACAAA----- 862
|||
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
|||
DB 862 ----- 862
|||
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
|||
DB 863 -----ACCTCAGACAGCAAGCAAGAG----- 883
|||
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
|||
DB 883 ----- 883
|||
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
|||
DB 884 GCCCACTAAGCTAGCG----- 901
|||
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySer 375
|||
DB 902 -----GAAATCAACAAGCCGAGGATC 925
|||
QY 376 MetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAla 395
|||
DB 926 CCCGAGAGTTCG----- 937
|||
QY 396 LysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAsp 415
|||
DB 938 -----AACGACGAC----- 946
|||
QY 416 GlyAlaGlyLysAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThr 435
|||
DB 946 ----- 946
|||
QY 436 GlyAlaGluLysAspIleValHisGlyHisArgGlySerAlaLeuSerMetArg 455
|||
DB 947 -----GTCGTTTCAGGTAAT 961
|||
QY 456 ArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGlyGlyGly 475
|||
DB 962 AAACAAAGC-----AACGTTGAAGGTGTA----- 985
|||
QY 476 TrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGly 495
|||
DB 985 ----- 985
|||
QY 496 PheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleVal 515
|||
DB 985 ----- 985
|||
QY 516 SerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeu 535
|||
DB 985 ----- 985
|||
QY 536 ValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAla 555
|||
DB 985 ----- 985
|||
```

QY 556 MetValHisProSerGluValAlaAlaLysGlySerArgTyrPheLysAspLeuPhe----- 573  
 Db 986 -----TGGAAAGAGCTTCTCTCTAT 1006  
 QY 574 ---GluProGlyValArgAlaLeuValGlyValGlyLeuGlnLeuGln 592  
 Db 1007 CCAAGCCCGCAATTCGCACATCGTAATCGCTGCCCTTGGTATTCACCTTCCACAA 1066  
 QY 593 PheAlaGlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluAlaGlyVal 612  
 Db 1067 GCGTCGGCGGTAGACCGCTGTTTGTACAGCCCGAGATCTTCCAAAGCGTGGGATT 1126  
 QY 613 AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuLeuSerSer 632  
 Db 1127 ACAAAACGACACATGAAGCTT-----CTTGCAACCGTGGCGCTTGGATT 1171  
 QY 633 LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer 652  
 Db 1172 GTTAAGACCGTGTTCATCTTGGCG-----GCTACGTTTACGTTTGGACCGCGTG 1219  
 QY 653 GlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672  
 Db 1220 GGTGCTGCTCGCTTATTGTCTAGTGTGGCGGCATGGTCTCGCTTCTCACGCTT 1279  
 QY 673 ValValSer---AsnLeuIleAspLeuGly-----ThrLeuAlaHisAlaLeuSer 689  
 Db 1280 GGCATCGCTCCTCATTTATTTGATCATTCGGAGAGGAATTAATGTGGCGCTTGGATCG 1339  
 QY 690 ThrValSer---ValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708  
 Db 1340 AGCATAGCATCGTGTGTGCTTACGTGGCCAGCTTCTCCATCGGTGGCGGTCCCATCAG 1399  
 QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728  
 Db 1400 TGGGCTATAGTCTTCTGAGATCTTCCCGTTGAGGCTGCGCGCGCARGGTGCGCGCGGGA 1459  
 QY 729 AlaPheThrPheThrIleGlyAspIleValThrTyrSerLeuProValMetLeuAsn 748  
 Db 1460 GTTGGCGTGAATAGGACCATAGCGCGTGTCTCAATGACTTTTCTGCTCCTCATCTAGA 1519  
 QY 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheVal 768  
 Db 1520 GCCATCACTATTGGTGAGCTTCTTCTTATTGTCGCAATTGCTACTGTGGGTGGATA 1579  
 QY 769 PheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr---Glu 787  
 Db 1580 TTTTWTACCGCTTGTGCTGAGACCCCGGGGAAACCGCTCGAGACATGGAAGGTCT 1639  
 QY 788 PhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800  
 Db 1640 TTTGGTACTTTTAGGTCCAAATCCAAACGACCAAGGCT 1678

RESULT 8  
 ; US-09-291-922-21  
 ; Sequence 21, Application US/09291922  
 ; Patent No. 6383776  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Steve  
 ; APPLICANT: Hitz, Bill  
 ; APPLICANT: Kinney, Tony  
 ; APPLICANT: Tingey, Scott  
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins  
 ; FILE REFERENCE: BB-1163  
 ; CURRENT APPLICATION NUMBER: US/09/291,922  
 ; CURRENT FILING DATE: 1999-04-14  
 ; EARLIER APPLICATION NUMBER: 60/083,044  
 ; EARLIER FILING DATE: April 24, 1998  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 21  
 ; LENGTH: 2017  
 ; TYPE: DNA  
 ; ORGANISM: Oryza sativa

US-09-291-922-21  
 Alignment Scores:  
 Pred. No.: 9,386-42 Length: 2017  
 Score: 494.50 Matches: 185  
 Percent Similarity: 35.71% Conservative: 101  
 Best Local Similarity: 23.10% Mismatches: 206  
 Query Match: 12.13% Indels: 309  
 DB: 22  
 Gaps: 22  
 US-10-051-909-32 (1-800) x US-09-291-922-21 (1-2017)  
 QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProAlaSerCysSerSerGln 43  
 Db 30 TTACTCTGACCCGACCTACTGTACACGCCCA---GAGCGAGCTCTCTCTCTCTGCA 86  
 QY 44 GluProValThrSerAspIleLeuGluAsp-----LysMetSerGly--- 58  
 Db 87 CCACCGAGAGTGTTCGCCGCGCTGCCGAGGCGGTGCCGCGAGAAAGAGGCAAC 146  
 QY 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75  
 Db 147 GTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTAC 206  
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
 Db 207 GATATCGGGGTGATGAGCGGGCGTGTGTGTACATCAAGAAGAGCTTCAATC---AGT 263  
 QY 96 GluProThrValGlu-----GlyLeuIleValSerMetSerLeuIleGlyAlaThr 112  
 Db 264 GACGGAGAGTGTGAGGTCTCATGGCATCTGMAACCTCTACTCGCTCATCGCTCTTC 323  
 QY 113 IleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIle 132  
 Db 324 GCG-----GCGGGCGGACGTGAGTGTGATCGGCGCGGTACACCATCGT 371  
 QY 133 LeuSerSerIleLeuTyrPhePheSerGlyLeuMetLeuTyrSerProAsnValTyr 152  
 Db 372 TTCGCGCGCTCATATCTTTCGCGGGGSGTTCCTCATGGGGTTCGCGCTCACTACGCC 431  
 QY 153 ValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuVal 172  
 Db 432 ATGCTCATTTCCGCGCGCTTCGTCGCGGATCGGCGGTGCGCTTCTGACGTCGTCGCG 491  
 QY 173 ProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuPro 192  
 Db 492 CCGGTGTACACCGCGGAGGTGTGCGCGCGTGTGCGCGGTGCGCTTCTGACGTCGTCGCG 551  
 QY 193 GlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeu 212  
 Db 552 GAGGTGTTCATCAACTTCGCGCATCTGCTCGGCTACGTCTCGAACTATGCTTTCTCCGCG 611  
 QY 213 SerPro---SerProAspThrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231  
 Db 612 TTCGCGCTGACCTCGGCTGCGCATCATGCTCGGCATCGCGCGCGCGCTCGTG--- 668  
 QY 232 PhePheGlyLeuThrIlePheTyrLeuProGluSerProAspTyrPheValSerLysGly 251  
 Db 669 CTGCTCGGCTCATGTGTCTCGCATCGCGAGTGTGCGCGGTGCTGCTCATGAGGGA 728  
 QY 252 ArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGly 271  
 Db 729 CGCTCGCGGACGCCAAGTGTGCTGGAGAGAC-----TCCGACACG--- 764  
 QY 272 GluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluTyr 291  
 Db 765 -----TCCGACACG----- 773  
 QY 292 IleIleGlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGln 311  
 Db 774 -----GCGGAGGAGCGCGCGGCGCTG----- 797  
 QY 312 IleThrLeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyPro 331

Db 797 ----- 797  
Qy 332 IleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSer 351  
Db 797 ----- 797  
Qy 352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGln 371  
Db 797 ----- 797  
Qy 372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391  
Db 797 ----- 797  
Qy 392 AspGlnHisAlaLysAsnGluGlnTrpAspGluAsnLeuHisArgAspAspGluGlu 411  
Db 797 ----- 797  
Qy 412 TyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSer 431  
Db 797 ----- 797  
Qy 432 ArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAla 451  
Db 797 ----- 797  
Qy 452 LeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAsp 471  
Db 798 -----GCCGAC 803  
Qy 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491  
Db 804 ATCAAGCCGCC----- 815  
Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511  
Db 816 -----GCCGCATCCCTGAGGAGCTCGAC 839  
Qy 512 GlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGluGlySerGluPhe 529  
Db 840 GCGCAGCTGTGACCGTCGCCCAAGAGAGGAGCGGAAC----- 878  
Qy 530 ValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPro 549  
Db 879 -----GAGAAG 884  
Qy 550 ArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrp 569  
Db 885 CCGGTG-----TGG 893  
Qy 570 LysAspLeuPhe-----GluProGlyValArgAlaLeuLeuValGlyValGly 586  
Db 894 AAGGAGCTCATCTGTCTCCCGACCCCGCCATCGCGGCATCTGTCTCGGGATCGGC 953  
Qy 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIle 606  
Db 954 ATCCACTTCTTCAGCATGCTGTGGGCATTCACCTCGCTCTTACAGCCCTCTCGTG 1013  
Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626  
Db 1014 TTCAAGAGCCCGGATTAACG-----AACGACAAACACTTCTTGGGCACCACTTGGCCG 1067  
Qy 627 SerIleLeuSerSerLeuThrThrLeuMetLeuProCysIleGlyPheAlaMet 646  
Db 1068 TTCGGTGTCAACCAAGAGCTTTTCATCTGTG-----GCCACT 1106  
Qy 647 LeuLeuMetAspLeuSerGlyArgPheLeuLeuLeuGlyThrIleProIleLeu 666  
Db 1107 TTCTTCATCAGCGCGTCGGCGCGCTGTGTGTCGGCAGACGCGCGGATATC 1166  
Qy 667 AlaSerLeuVal-----IleLeuValValSerAsnLeuIleAspLeuGly 681

Db 1167 CTCTCCCTCATCGCCTCGCGCGGGCTCACCGTCGTGGCCAGCACCCCGGCCAAG 1226  
Qy 682 ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701  
Db 1227 ATACCTTGGGCCATCGCGCTTAAGCATCGCCTCCACCCCTCGCTACGTCGCTTCTTCTCC 1286  
Qy 702 MetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArg 721  
Db 1287 ATCGGCTTGGCCCATCAGCTGGGTGTACAGCTCGGAGATCTTCCGCTCCAGGTGGC 1346  
Qy 722 -----GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739  
Db 1347 GCGCTGGGCTGCTCGCTCGCGCTCGCGCCACCGGCTACCAGCGGCGTCATCTCCATG 1406  
Qy 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759  
Db 1407 ACCTTCTCTG-----TCGGTGTCCAAGGCCATCACCATCGCGCGCAGCTTCTTCTCTAC 1460  
Qy 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779  
Db 1461 TCGGCGATCGCGCGCTCGCTGGGTGTCTTACACCTTACCTCCGCGAGACCCGCGGC 1520  
Qy 780 MetProLeuGluValIleThrGluPhePheAlaValAlaLysGlnAlaAlaLys 799  
Db 1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTT-----GGCGACACGCGCGCTCGGAA 1574  
Qy 800 Ala 800  
Db 1575 TCA 1577

RESULT 9  
US-09-291-922-19  
; Sequence 19, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 19  
; LENGTH: 1914  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-291-922-19

Alignment Scores:  
Pred. No.: 1,78e-41 Length: 1914  
Score: 491.50 Matches: 172  
Percent Similarity: 35.36% Conservative: 107  
Best Local Similarity: 21.80% Mismatches: 201  
Query Match: 12.06% Indels: 309  
DB: 3 Gaps: 18

US-10-051-909-32 (1-800) x US-09-291-922-19 (1-1914)  
Qy 41 SerSerGlnGluProValThr-----SerAspAspIle-----LeuGluAspLys 55  
Db 35 TCAAGAAGTAGCGTTACGATGGCTTCGACGAGCTGCCAAGGCCGTCGAGCCCGCAGG 94  
Qy 56 MetSerGly-----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeu 71  
Db 95 AAGAAGGGCAACGTCGAAGTATGCTTCATATGTGCCATCTCGGCTCCATCGGCTCTGTG 154  
Qy 72 LeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPhe 91

155 ATCCCTGGCTATGACATTTGGGTGATGAGTGGAGCGGCATGTATACATCAAGAGGACCTG 214  
 QY  
 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111  
 Db  
 215 AATATC---ACGACGTGCGAGTGGAGTCTGTATCGGGATCCTCAGTCTCTAC----- 265  
 QY  
 112 ThrIleValThrPheSerGlyPro---LeuSerAspSerIleGlyArgAspProMet 130  
 Db  
 266 TCGCTGTTCCGATCCTTCTGCTGGCGCGGAGCTCCGACAGATCGCGCGCGCTTACCC 325  
 QY  
 131 LeuIleLeuSerSerIleLeuThrPhePheSerGlyLeuIleMetLeuThrSerProAsn 150  
 Db  
 326 GTGCTGTTCCGCGCTGTCATCTCTCTGCGGCTCGTTGCTCATGGGTTCGCGGTCAAC 385  
 QY  
 151 ValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170  
 Db  
 386 TAGCGCATCTCATGGCGGCGCTCTGTTGGCGGAGTCGGTGTGGGTACGGGGGCATG 445  
 QY  
 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190  
 Db  
 446 ATCGCGCCGTATACAGCGCGGAGATCTCGCTCGCGGCTCGCGTGTCTGTACCAACC 505  
 QY  
 191 LeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMet 210  
 Db  
 506 TTCGCCGAGGTTTCATCAACATCGGCATCTGCTGGCTACCTGCCAATTCGCGTTC 565  
 QY  
 211 SerLeuSerPro---SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSer 229  
 Db  
 566 GCGCGCTCCGCTCCACCTCGGCTGGCGGTCATGCTCGCATTCGCGGCGGATTCGCTCC 625  
 QY  
 230 LeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSer 249  
 Db  
 626 ---GGCTGTCTCGCTCTCTGTTGTTCTGATGCCGAGTCGCCGTGGTGTGCTCTG 682  
 QY  
 250 LysGlyArgMetAlaGluAlaIleLysValLeuGlnLysLeuArgGlyLysAspVal 269  
 Db  
 683 AAGGCGCGCTCGCGACGCGAGGCTGTGCTAGAGAAG----- 721  
 QY  
 270 SerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGlyAspThrSerIleGlu 289  
 Db  
 721 ----- 721  
 QY  
 290 GluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLys 309  
 Db  
 721 ----- 721  
 QY  
 310 GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329  
 Db  
 722 -----ACCTCTGCCAGCGCCAGAGAGCGCGCGAGGCTGCGC----- 760  
 QY  
 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349  
 Db  
 760 ----- 760  
 QY  
 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369  
 Db  
 760 ----- 760  
 QY  
 370 ProGlnAlaGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389  
 Db  
 760 ----- 760  
 QY  
 390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp 409  
 Db  
 760 ----- 760  
 QY  
 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu 429  
 Db  
 760 ----- 760  
 QY  
 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly 449  
 Db  
 760 ----- 760

450 SerAlaLeuSerMetArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSer 469  
 Db  
 760 ----- 760  
 QY  
 470 ThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGluAsn 489  
 Db  
 761 ---GACATCAAGCGCGC----- 775  
 QY  
 490 GlyArgLysGluGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509  
 Db  
 776 ---GCGGGGATTTCCGAAGGCG 793  
 QY  
 510 ArgArgGlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGlu 525  
 Db  
 794 CTCGACGGGAGCTAGTACCTGACCTACCGCAAGGAGGAGCGCGGTGAGTTG----- 847  
 QY  
 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545  
 Db  
 847 ----- 847  
 QY  
 546 LeuAlaGluProArgMetSerAspAlaMetValHisProSerGluValAlaAlaLys 565  
 Db  
 847 ----- 847  
 QY  
 566 GlySerArgTrpLysAspLeuPhe-----GluProGlyValArgArgAlaLeuLeu 582  
 Db  
 848 ---CAGGTGTGAAGAAGCTCATCTCTCCCGACCCCGGTGTCTCCGACGATCTGCTC 904  
 QY  
 583 ValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyr 602  
 Db  
 905 TCGCGCGTGGTCTCCACTTCTTCAGCAGGCTTCTGGCAGCGACTCCGTCTGTCAGTAC 964  
 QY  
 603 ThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSer 622  
 Db  
 965 AGCGCGCGCTGTTCAGAGCGCGGGGATCCCGACGACCAAGCTCTCTGGCGCTCAC 1024  
 QY  
 623 SerAlaSerAlaSerIleLeuIleSerSerLeuThrLeuLeuMetLeuProCysIle 642  
 Db  
 1025 TGGCGG-----GTGGCGGTGACCAAGCGTTCTTCTCATCTG----- 1060  
 QY  
 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle 662  
 Db  
 1061 ---GTGGCCACGTCTCTGCTGGACCGCGCGGCTGCGCTCTGCTGTGATCAGCAG 1177  
 QY  
 663 ProIleLeuIleAlaSerLeuValIleLeu-----ValValSerAsnLeu 677  
 Db  
 1118 GCGGGATGATGATGCTCTGCTCATCTGCTCGGGTCTGGGCTCACCGTCCGGGGCATCNC 1177  
 QY  
 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697  
 Db  
 1178 CCGACACCAAGTCTCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1237  
 QY  
 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717  
 Db  
 1238 GCCTTCTCTCCATCGGCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297  
 QY  
 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737  
 Db  
 1298 CTGAGTGTGGCGCTGGGCTTCTGCGGTGGGTGCGGAGCAACCGCTCACCGCGCGC 1357  
 QY  
 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757  
 Db  
 1358 GTCATCTCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417  
 QY  
 758 IleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777  
 Db  
 1418 CTCTACTCCCGCATCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1477  
 QY  
 778 LysGlyMetProLeuGluValIleThrGluPhePhe-----Ala 790  
 Db  
 1478 CCGCGCGGCGCTGGAGGAGATGGGCAAGCTGTTCGGCATGCGACGACGCGGATGGCT 1537

```
QY 791 ValGlyAlaLysGlnAlaAlaLys 799
Db 1538 GAAGAAGCAGACGCGCGCAAG 1564

RESULT 10
US-09-291-922-27
; Sequence 27, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER APPLICATION NUMBER: 1999-04-14
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-27

Alignment Scores:
Pred. No.: 3 15e-41 Length: 1872
Score: 489.00 Matches: 165
Percent Similarity: 34.22% Conservative: 106
Best Local Similarity: 20.83% Mismatches: 215
Query Match: 12.00% Indels: 306
DB: 3 Gaps: 14

US-10-051-909-32 (1-800) x US-09-291-922-27 (1-1872)
QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCysSerGln 43
Db 119 ATGGCTTCTGCGCTCCCGGAGCGGGGCGAGTCATCCAGGAAC----- 166
QY 44 GluProValThrSerAspAlaLeuGluAspLysMetSerGlyAlaValLeuValAla 63
Db 167 -----NAGGGCAATTTCAAGTACGCTTCACCTCGGCC 199
QY 64 IleValAlaSerIleGlyAsnLeuLeuGlnGlyTyrAspAsnAlaThrIleAlaAla 83
Db 200 CTCTGTCTTCAATGCCACCATCGCTCGGTACGACCTTGGGTGATGACGGTGCG 259
QY 84 ValLeuTyrIleLysGluPheGlnLeuGlnAsnGlu-----ProThrValGluGly 101
Db 260 TCGCTGTACATCAAGAGGACCTGCAGATCAGCGAGCTGCGATGATGATGCGGC 319
QY 102 LeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeu 121
Db 320 ATCTCAGCGGTGACGCGCTCATCGGTCTCTCTC-----GCCGCGAGGACG 367
QY 122 SerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyrPheSer 141
Db 368 TCGATGCGGTGCGCGCGCGCTACCGCTCTTCGCGCGCCCATCTTCAACACGCGC 427
QY 142 GlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeuLeuAlaArgPheValAsp 161
Db 428 TCCTTGCTCATGGGCTTCGGGTCACTACGCCATGCTCATGTCGGCGCTTCGTACC 487
QY 162 GlyPheGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPro 181
Db 488 GGAATCGCGTGGCTACGCCATCGTGTCCGCGCAGTGTACACGCCCGAGGTGTCGCC 547
QY 182 SerGluIleArgGlyLeuAnThrLeuProGlnPheSerGlySerGlyMetPhe 201
Db 548 GCCTCGCGCGCGGCTCTCTCAGCTCTTCCACCGAGGTGTTCATCAATGTGGGCATCCTC 607

202 LeuSerTyrCysMetValPheGlyMetSerLeuSerPro---SerProAspTrpArgIle 220
608 CTTGGCTACGCTTCCAACTACGCTTCGCGCGCCTCCCGCTCCACCTCAGCTGCGCGTC 667
221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
668 ATGCTCGGATCGCGCGCTCCGTCGCGC---CTGCTGCGCTCATGTTGCGCATG 724
241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysValLeu 260
725 CCGAGTCTCTCTGCTGCTGCTCATGAAGGCGCGCTCGCGAGCCGAGGCGGCTCTG 784
261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
784 ----- 784
281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
784 ----- 784
301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
784 ----- 784
321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
784 ----- 784
341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetMetAspProIleValThr 360
784 ----- 784
361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
784 ----- 784
381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
785 -----GCCAAGACCTCCGACACG 802
401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp 420
803 CCGAGAGCGCGCTGAGCGCTTACCAGATCAGCTGCC----- 844
421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
844 ----- 844
441 AspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
844 ----- 844
461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480
844 ----- 844
481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
844 ----- 844
501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
845 -----GCCGCGATCCCTAGGGAACCTGACGCGACGCTGCTCATGCTCAAG--- 892
521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuLeuGlnSerAla 540
892 ----- 892
541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
893 -----ACAAAGCGCGCGCAGGAGCAGGTG----- 919
561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPhe-----GluProGlyVal 577
```

```
Db 920 -----TGAAGAGGCTCATCTTTTCGCCGACCCAGCCATG 955
QY 578 ArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsn 597
Db 956 CGGCGCATACTGCTCGCGCGCTCGCATCTTTCAGAGCGGCGGCTCCGAC 1015
QY 598 GlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSer 617
Db 1016 TCGCTGCTGCTATAGCCACCGCTTCCAGCGCGGCATACCGCGCACACAC 1075
QY 618 LysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrLeuLeu 637
Db 1076 CTGCTCGGC-----CCACATCGCCATGGGGGTGATCAGACAGCTCTTC 1120
QY 638 MetLeuProCysIleGlyPheAlaMetLeuMetAspLeuSerGlyArgArgPheLeu 657
Db 1121 ATCTGTG-----GTGGCCACGTTCCAGCTCGACCGCTCGGCGGCGCGCTG 1168
QY 658 LeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal-----IleLeu 672
Db 1169 CTGCTGACCAGCAGCGCGCATGCTCGCTGTCTCATCGGCCTCGGACGGGCTCACC 1228
QY 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
Db 1229 GTCTGGGTGCGCACCGCGCAAGCTCCGTCGGGCCATCGGCCTGTGATCGTGTCC 1288
QY 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 1289 ATCTTGGCTAGTGTCTTCTTCATCGCCCTCGGCGCCCTCACCGAGGTGATCACC 1348
QY 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
Db 1349 TCGAGAGTCTTCCCACTCGCGGTGCGCGCTGCGCTTCGGCTGGGACGTCATGCAAC 1408
QY 733 TrpIleGlyAspIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db 1409 CGCGTCACCGCGCGCGCTCTCCATGTCTTCTGTCTTGTCTTGTCCAAAGGCCATCACC 1468
QY 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db 1469 GCGCGCAGCTTCTTCTGTAGCGCGCATCGCGCGATAGGATGATTTTCTTCTTACC 1528
QY 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe----- 789
Db 1529 TTCATTCGGAGACGCGTGGCTCGCTCGAGGAGATAGGAAGCTTTTCGGCATGACG 1588
QY 790 -----AlaValGlyAlaValGlnAlaAlaAla 799
Db 1589 GACACGGCGCTCGAAGCCACAGACACCGCCACGAAA 1624
```

## RESULT 11

```
US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER FILING DATE: 60/083,044
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-25
```

## Alignment Scores:

```
Pred. No.: 1,948-38 Length: 2089
Score: 463.50 Matches: 166
Percent Similarity: 35.82% Conservative: 108
Best Local Similarity: 21.70% Mismatches: 200
Query Match: 11.37% Indels: 291
DB: Gaps: 17
```

US-10-051-909-32 (1-800) x US-09-291-922-25 (1-2089)

```
QY 52 LeuGluAspLysMetSerGly-----AlaValLeuValAlaIleValAlaSer 67
Db 104 GTTCAGCCCAAGAGGCAACGTGAGGTTCGCTTCGCTCGCATCTCGCTCC 163
QY 68 IleGlyAsnLeuLeuGlnGlyTyrAspAsnAlaThrIleAlaAlaValLeuTyrIle 87
Db 164 ATGACCTTCATCTCTCGCTAGACATCGCGGTGATGAGCGGAGCGTCTGTATC 223
QY 88 LysLysGluPheGlnLeuGlnAsnGluProThrValGlu-----GlyLeuIleVal 104
Db 224 CAGAAGGATCTGAGATC---AACGACACCAGCTGGAGGTCTCATGGGCATCTCAAC 280
QY 105 SerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSer 124
Db 281 GGTACTCGCTATTGGCTCTTCGCG-----GCGGGCGGACGTCGACTGG 328
QY 125 IleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIle 144
Db 329 ATCGCGCGCGGTTCACCATGCTTCGCGCGCTCATCTTCTTCGCGGCGCCTCATC 388
QY 145 MetLeuTyrSerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGly 164
Db 389 ATGGGCTTCTCGCTCACTACCGCATGCTCATGTCGGCGCCTTCGTGCGCGCATCGC 448
QY 165 IleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIle 184
Db 449 GTGGGATAGCTCTCATATCGCGCGCTGAAACACGCGGAGGTGTCCCGCGCTGTGCC 508
QY 185 ArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyr 204
Db 509 CGTGGGTCTCATCTCTCCGAGGTGTTCATCAACTTCGCGCATCTCTCGGATAT 568
QY 205 CysMetValPheGlyMetSer---LeuSerProSerProAspTyrArgIleMetLeuGly 223
Db 569 GTCTCAAATTCGCTTCGCGCGCTCTCTCCCTCGCGCTCGCGCATATATCTCGGC 628
QY 224 ValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSer 243
Db 629 ATAGCGCGGTGCTCGCTCGCTC---CTGCTCGGTTCATGGTCTCGCATCCCGAGTCT 685
QY 244 ProArgTyrLeuValSerLysGlyArgMetAlaGluAlaLysValLeuGlnLysLeu 263
Db 686 CCGCGTGGCTCTCATGAGGCGCGCTCTCGCGAGCCCAAGGTGTGCTTGCACAGACG 745
QY 264 ArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGly 283
Db 745 ----- 745
QY 284 GlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspLeu 303
Db 746 TCCGACACG----- 754
QY 304 ValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTyr 323
Db 755 -----CCGGAAGAGAGCGCGCGAGCGC 775
QY 324 IleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArg 343
Db 776 ATCGCC----- 781
QY 344 HisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGly 363
```

781	----	781
Db		
364	SerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsn	383
Qy		
781	----	781
Db		
384	PheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGlu	403
Qy		
781	----	781
Db		
404	AsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp	423
Qy		
782	-----GACATTAGACTGCCGCC-----	799
Db		
424	AsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleVal	443
Qy		
799	----	799
Db		
444	HisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGly	463
Qy		
799	----	799
Db		
464	GlyAspGlyValSerSerThrAspIleGlyGlyTyrTrpGlnLeuAlaTrpLysTrpSer	483
Qy		
799	----	799
Db		
484	GluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGln	503
Qy		
799	----	799
Db		
504	GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspVal	523
Qy		
800	---GGCATCCCTCTGGGCCTCGACGGCAGCTGTGCTCCCGTCCGCAAA-----	844
Db		
524	PheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSer	543
Qy		
845	-----AAC847-----	
Db		
544	LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla	563
Qy		
848	AAAGAGACGACGAGGAGAGCGGCTTTGNAGGACCTCATCTCTGTCACCGACCATAGCC	907
Db		
564	AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuVal	583
Qy		
908	-----ATGCCCCACATCTCTCATCGCG928-----	
Db		
584	GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThr	603
Qy		
929	GGAAATCGCATCTCACTTCTTCCACACAGCTCTTCGGGCATCGACGCGCTGCTCTACAGC	988
Db		
604	ProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer	623
Qy		
989	CCGCTAGTTTCAGAGCGCGGCATCAGC-----GGCGACAGCCGT1030-----	
Db		
624	---AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle	642
Qy		
1031	CTCCGCGCACCCACCGTGGCGGTGCGGGCCACCAATAACGCTCTTCATCTCTG-----	1081
Db		
643	GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIle	662
Qy		
1082	---GTGGCCACCTTCTCTCTCGACCGCATCGCGCGCGCGCTGTGTGCTACACGACGC	1138
Db		
663	ProIleLeuIleAlaSerLeuVal-----IleLeuValValSerAsnLeu677-----	
Qy		
1139	GGCGGCATGCTGTGCTCTTAGTGGGCTGCGAGCGGGGCTCACCGTCATCAGCGGCCAC	1198
Db		
678	IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe	697
Qy		
1199	CCGACAGCAAGATCACCTGGGCCCATCGTCTCTGTGCATCTTCTGCATCATGGCCTACGTG	1258
Db		
698	CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro	717
Qy		
1259	GCCTTCTTCTCCATCTCGGCTCGGCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG	1318
Db		

Qy	718	ThrArgValArgGlyLeuCySIIeAlaIleCysAlaphethrPhetrlieGlyAspIle	737
Dd	1319	CTGCACGTGGCGCGTGGCTCTCCCTGGGCGTGCCGTCAACCGCCTGACCAAGCGGC	1378
Qy	738	IleValThrTyrrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer	757
Dd	1379	GTGATCTCCATGACCCTCATTTTCGTGTCCAAGGCCATGACCATCGGCGGCGCTTCTTC	1438
Qy	758	IleTyzAlaValValCysLeuIleSerPheValPheValPheLeuIlysValProGluThr	777
Dd	1439	CTCTTCGGCGGCATCGCCTCATTCGCATGGTGTCTTCTTCGCTACCTGCGGAGACC	1498
Qy	778	LysGlyMetProLeuGluValIleThrGlupPhePheAlaValGlyAla-----LysGln	795
Dd	1499	CGCGCGCGCACGCTGGAGGACATGAGTCTGCTTTCGGCAACCGGCCACGCAACAGCAG	1558
Qy	796	AlaAlaAlaLysAla 800	
Dd	1559	GCGCGCGCGGAAGCC 1573	
<b>RESULT 12</b>			
US-09-489-039A-4762			
; Sequence 4762, Application US/09489039A			
; Patent No. 6610836			
; GENERAL INFORMATION:			
; APPLICANT: Gary Breton et. al			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLBES			
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 2709.2004001			
; CURRENT FILING DATE: US/09/489, 039A			
; PRIOR APPLICATION NUMBER: 2000-01-27			
; PRIOR FILING DATE: 1999-01-29			
; NUMBER OF SEQ ID NOS: 14342			
; SEQ ID NO 4762			
; LENGTH: 1431			
; TYPE: DNA			
; ORGANISM: Klebsiella pneumoniae			
US-09-489-039A-4762			
<b>Alignment Scores:</b>			
Pred. No.:	5,62e-37	Length:	1431
Score:	447.00	Matches:	158
Percent Similarity:	33.68%	Conservative:	98
Best Local Similarity:	20.79%	Mismatches:	186
Query Match:	10.97%	Indels:	318
DB:	4	Gaps:	18
US-10-051-909-32 (1-800) x US-09-489-039A-4762 (1-1431)			
Qy	42	SerGlnGluProValThrSr-----AppAspIleLeuGluAsp---LysMetSerGly---	58
Dd	4	TGCACAGAATCAATAACTCAACTGGAGGGCGTTATGCTTGACAAACAAAAACAAGGCGGT	63
Qy	59	-----AlaValLeuValAlaIleValAlaSerIleGlyAenLeuLeuGln	73
Dd	64	TCGACACAGACTATGAGTCTTCGTCTGTTTTCTCGCGCGGTGGCTGGCTGCTGTTCT	123
Qy	74	GlyTrpAspAsnAlaThrIleAlaAlaValLeuTyriIleLysIleValPheGlnLeu	93
Dd	124	GGCCTTGATATCGTGTATTTCGGGTGCTTACCTCCCTTTATTCGCAATGAGTTCCAGATT	183
Qy	94	GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle	113
Dd	184	TCGCCCCAC-----ACCCAGGAGTGGTGGTTCAGCTCCATGATGTCGGGCGCTCGGTC	237
Qy	114	ValThrThrPheSerGlyProLeuSerAspSerIleGlyVargArgProMetLeuIleLeu	133
Dd	238	GGCGCGGTGGCAGCGGCTGGCTCTCTTTCAAACCTGGCGCGGAAAAGACCTGATGATC	297
Qy	134	SerSerIleLeuTyrrPhePheSerGlyLeuIleMetLeuTrpSerProAenValTyrrVal	153

Alignment Scores:		
Pred. No.:	5.62e-37	Length:
Score:	447.00	Matches:
Percent Similarity:	30.6%	Conservative:
Best Local Similarity:	23.78%	Mismatches:
Query Match:	10.97%	Indels:
DB:	4	Gaps:
		18
		1431

US-10-051-909-32 (1-800) x US-09-489-039A-4762 (1-1431)

Qy	42	SerGlnGluProValThrSer-----AspAspIleLeuGluAsp---lysMetSorGly---	58
		:    :    :    :    :    :	
Db	4	TCGACAGAATCAATAACTCAACTGGAGGGCGTTATGCTGCACACAAAAACAAGGGCGT	63
Qy	59	-----AlaValLeuValAlaIleValAlaSerIleGlyAenLeuGln	73
		:    :    :    :    :    :	
Db	64	TCGACACAAGACTATGACGTTCTTCGTGTCGTTCCTCGCCGCGTGGCTGCCTGCTGTTTC	123
Qy	74	GlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleValysGluPheGlnLeu	93
		:    :    :    :    :    :	
Db	124	GGCCTTGATATCGGTGTTATGCGGGTGCCTTACCCTTTATTCCCAATGAGTTCAGATT	183
Qy	94	GlnAsnGluProThrValGluGluLeuIleValSerMetSerLeuIleGlyAlaThrIle	113
		:    :    :    :    :    :	
Db	184	TCGCGCCCAAC-----ACCCAGAGTAGTGGGTGCAGCTCCATGATGTTCGGGGCTGGCGCT	237
Qy	114	ValThrThrPheSerGlyProLeuSerAspSerIleGlyValArgProMetLeuIleLeu	133
		:    :    :    :    :    :	
Db	238	GGCGCGGTGGCAGCGGTGGCTCTCTTTCAAACTGGCGCGGAAAAAGAGCCCTGATGATC	297
Qy	134	SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal	153
		:    :    :    :    :    :	





TYPE: DNA			
ORGANISM: Klebsiella pneumoniae			
US-09-489-039A-4731			
Alignment Scores:			
Prod. No.:	3 15e-36	Length:	1545
Score:	440.50	Matches:	164
Percent Similarity:	34.05%	Conservative:	105
Best Local Similarity:	20.76%	Mismatches:	206
Query Match:	10.81%	Indels:	315
DB:	4	Gaps:	16
US-10-051-909-32 (1-800) x US-09-489-039A-4731 (1-1545)			
QY	25	ProSerValValLeuAlaLeuProGlyProLeuProAlaSer	39
DB	37	CCATATCTTTGGCTTATCTCTGCTGACCCATTCATCCACGACGTCAGTACTTTTCGTCTTA	96
QY	40	-----CysSerSerGlnGlu-----ProValThrSerAspAspIleLeuGlu	53
DB	97	TGCTTTACTCTGTGTGGCAGGAAAAAATGACTTCAATCAGTAAACGACTCTACATTATCG	156
QY	54	AspLysMetSerGlyAlaVal-----LeuValAlaIleValAlaSerIle	68
DB	157	CCGGGACGCAAGCTGATACCCCGGATCACTGTTTTCATCCGCGCGGGTG	216
QY	69	GlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrlleLys	88
DB	217	GCTGGCTTGCTTTGGCTTGGATATCGCGGTGATATCCGGAGCGTTGGCCCTTTATAACC	276
QY	89	LysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu	108
DB	277	GACCATTTTCACTTATCCAGCCAG-----CTTCAGGAGTGGGTGTAGCAGCATGATG	330
QY	109	IleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArg	128
DB	331	TTGGGGCGCGGATAGCGCGCTGTTTAAACGGTGGCTGTCTTCGCGCTTGGCCGTA	390
QY	129	ProMetLeuIleLeuSerSerIleLeuTyrllePheSerGlyLeuIleMetLeuTrpSer	148
DB	391	TACAGCTGATGCGGGGCGGTACTCTTTGTGCGCGCTCTATCGATCCGCTTTTGCC	450
QY	149	ProAsnValTyrlleValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAla	168
DB	451	GCCAGCGTGGAGTGTCTGTGGCGCGCGGTGTGTGGCGTGGCAGTGGCGGATTGCC	510
QY	169	ValThrLeuValProLeuTyrlleSerGluIleAlaProSerGluIleArgGlyLeuLeu	188
DB	511	TCTTATACCGCGCGCTGTACCTCTCCGATGCGCCAGCAGACGTCGCGGAAATG	570
QY	189	AsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrlleCysMetValPhe	208
DB	571	ATCAGTATGATATCAGCTGATGGTCACTTGTGGCATTTGTGGCGTT-----CTTTCC	624
QY	209	GlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIlePro	228
DB	625	GATACCGCTTTAGTACAGCGTAACGTCGCGGCCATGCTGGCGGTGTGGCGCTGCCG	684
QY	229	SerLeuPhePheGlyLeuThrIlePheTyrlleProGluSerProArgTrpLeuVal	248
DB	685	CGCGTATCTGTATCATCTTGTGCTCTTT-----TTGCCGAACAGCCGCGCTGGCGG	741
QY	249	SerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAsp	268
DB	742	GAGAGGGAGCCATATCGAAGCGGAAGATGCTCGGATGCTGGC-----789	
QY	269	ValSerGlyGluLeuSerLeuLeuGluGlyLeuValGlyGlyAspThrSerIle	288
DB	790	-----GATACCTCG-----798	
QY	289	GluGluTyrlleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAsp	308
DB	799	-----GAAAGCGCGCGACGACTT-----819	

QY	309	LysGluGlnIleThrLeuTyrlleGlyProGluGluGlyGlnSerTrpIleAlaArgProSer	328
DB	819	-----819	
QY	329	LysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetVal	348
DB	820	-----AACGAGATC	828
QY	349	AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn	368
DB	829	CGTGAGAGCTCAAGCTGAAGCAG-----852	
QY	369	MetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPhe	388
DB	852	-----852	
QY	389	SerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluAlaAsnLeuHisArgAsp	408
DB	852	-----852	
QY	409	AspGluGluTyrlleAlaSerAspGlyAlaGlyGlyAspTyrlleGluAspAsnLeuHisPro	428
DB	852	-----852	
QY	429	LeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArg	448
DB	852	-----852	
QY	449	GlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSer	468
DB	852	-----852	
QY	469	SerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGlu	488
DB	853	-----GGCGGTTGG-----861	
QY	489	AsnGlyArgLysGluGlyGlyPheLysArgValTyrlleHisGlnGluGlyValProGly	508
DB	861	-----861	
QY	509	SerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGlu	528
DB	861	-----861	
QY	529	PheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGlu	548
DB	862	-----GCCTTGTTTAAG-----873	
QY	549	ProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArg	568
DB	873	-----873	
QY	569	TrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGln	588
DB	874	-----ATCAATCGTAACGTCGCGCGCGGTGTCTCGGCAATGCTGTCGAC	921
QY	589	IleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrlleThrProGlnIleLeuGlu	608
DB	922	CGCATCGCAATTCACCGCATGACATCATCTACTATGCGCGCGGTATCTTTAA	981
QY	609	GlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIle	628
DB	982	ATGGCG-----GGCTTTACCACTACTGAACAGCATG	1014
QY	629	LeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeu	648
DB	1015	ATCGCACCCCTGTGTGGCGCTTATGTTTTCGCCACCTTTATGCGGTGTACACA	1074
QY	649	MetAspLeuSerGlyArgArg-----PheLeuLeuGlyThrIleProIleLeuAla	667
DB	1075	GTGGATAAGCGCGCGCGCAAGCCAGCGCTGAATAATCGCTTTAGCGTGATGGCGCTGGC	1134

Qy	668	SerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeu	687
	:::     :::		:::     :::
Db	1135	ACTCTGGTGGCTACTGCTGATGCAGTTCGACAATGGCACC---GCATCCAGCGCG	1191
	:::     :::		:::     :::
Qy	688	LeuSerThrGlyValIleValTyPheCysCys-----PheValMetGlyPhe	704
	:::     :::		:::     :::
Db	1192	CTCTCTGGGCTCTCCGTCGCATGACCATGATGTATTGCCGGGTATCGATGAGCGCG	1251
	:::     :::		:::     :::
Qy	705	GlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValargGlyLeuCys	724
	:::     :::		:::     :::
Db	1252	GCGCCGGTGTGTGGATCTCTGCTCCGATCCAGCCGCTAAATTCGCCGCACTTCGGT	1311
	:::     :::		:::     :::
Qy	725	IleAlaIleCysAlaPheThrPheIleGlyAspIleIleValThrTySerLeuPro	744
	:::     :::		:::     :::
Db	1312	ATCACCTGCTGCACCACCACCACCTGGGTGTGCAACATGATCATCGCGGCCACTTCTTCCTG	1371
	:::     :::		:::     :::
Qy	745	ValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleIleTyAlaValValCysLeu	764
	:::     :::		:::     :::
Db	1372	ACGCTGCTTGACGGGATGGCGCCGCCGCCACCTTCTGGCTCTACACGGCGCTCAACGTG	1431
	:::     :::		:::     :::
Qy	765	IleSerPheValPheValPheLeuIleValProGluThrIysGlyMetProLeuGluVal	784
	:::     :::		:::     :::
Db	1432	GCCTTTATCGGCATCACCTTCTGGCTGATCCCGGNAACCAAGATGTACACCTCGAGCAC	1491
	:::     :::		:::     :::
Qy	785	IleThrGluPhePheAlaValGlyAlaLys	794
	:::     :::		:::     :::
Db	1492	ATTGAGCGCAACCTGATGGCGGGCGAGAAG	1521
	:::     :::		:::     :::

```

RESULT 14
US-09-489-039A-4560
; Sequence 4560, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4560
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4560

```

Alignment Scores:	8.15e-35	Length:	1506
Pred. No.:	427.00	Matches:	155
Score:	33.29%	Conservatives:	104
Percent Similarity:	19.92%	Mismatches:	201
Best Local Similarity:	10.48%	Indels:	318
Query Match:	4	Gaps:	17
DR:			

US-10-051-909-32 (1-800) X US-09-489-039A-4560 (1-1506)

Qy	33	GlyProLeuProAlaSerCysSerSerGlnGluPro-----ValThrSer	48
Db	19	GGCCCG--CCTGAACCTACTGTTCCCTCGTTCCTTATAGAGGAATCATTTATGAAC	75
Qy	49	AspAspIleLeuGluAspLysMetSerGlyAlaValLeuValaIleValaSerIle	68
Db	76	AACGCGCAGACACATCTGAAAAATGGCGCTACGTCGTGGACGAATTTGTGTGTGCCGCCCTGC	135
Qy	69	GlyAsnLeuLeuGlnGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrllys	88
Db	136	GGTGGTTTACTGTTTGGCTATGACTGGGTGGTGATTGGCGGGCGCTAAGGCATTTTATGA	195
Qy	89	LysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu	108
Db	196	GCCTGTGTTTCAATT---ACGGACCCCGGGCGACATCCGGCTGGCGCATGAGCTCAGGGCTG	252

QY	109	IleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArg	128
DB	253	TTGGCGTGTATTTCGCGCGCATTAATTTTCGGATGTCGCGACACAAATGGGGCGCAAG	312
QY	129	ProMetLeuIleLeuSerSerIleLeuTyThrPhePheSerGlyLeuIleMetLeuTrpSer	148
DB	313	CTGGCATTAATTTTCGCGCGTCTGTTTCAGCGCTCGCGCTGGCGGAGCGGCTCGCC	372
QY	149	ProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAla	168
DB	373	AGTCATTTTCGATATGTTTGGTGTATACCGCATTTGCGCGCGTAGGATGTCCTGGCT	432
QY	169	ValThrLeuValProLeuTyTrpIleSerGluIleAlaProSerGluIleArgGlyLeuLeu	188
DB	433	TCCGCTCTCAGCCGCTTATCATTTGCCGATGTCAGCCCGCAGAGAAAGAGACGTTTT	492
QY	189	AsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPhe	208
DB	493	GTCGGCGTCAATCAGTCACCATCGTATTTGGCTGTCGCGCTCAAGTAAATCAATCTG	552
QY	209	GlyMetSerLeuSerProSerPro	216
DB	553	ATGATTGCTGAACCGTGGAGCCGGGGCGCAGCAGCATGATTGGACAGCTGGAAT	612
QY	217	-----AspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe	233
DB	613	GGGCAGATGGGCTGGCTGGATGTTTCGTCGGAACTGGTGCCGCGCACTGGCGTTCTG	672
QY	234	GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet	253
DB	673	GTCCCTG---ATGTTTTTTGTCCTCCGAGTCGCGCGCTGCATGAAGCCGCTAAACCG	729
QY	254	AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu	273
DB	730	GAGCGCGCGCGCTGCGCTGGAACGC	756
QY	274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle	293
DB	757	-----ATT	759
QY	294	GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr	313
DB	760	GGTTCGGCC	768
QY	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet	333
DB	768	-----	768
QY	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
DB	768	-----	768
QY	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
DB	768	-----	768
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
DB	768	-----	768
QY	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluTyrAla	413
DB	768	-----	768
QY	414	SerAspGlyAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
DB	769	-----GACTATGCCGACAGGATCCTG	789
QY	434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer	453
DB	790	-----CGTAAATCGCCGAT	804

```
QY 454 MetArgGlnThrLeuLeuGlyGlyAspGlyValSerSerThrAspIleGly 473
DB 805 -----ACCCTG-----
QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyAlaGlyAspGlyGlu 493
DB 811 -----GAAAGGATACAAATAA-----
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
DB 828 -----
QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
DB 828 -----
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
DB 829 -----GTCTCTACCGCGCGCTGTGGCT-----
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
DB 852 -----
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
DB 853 ---CCCAGGTGAACCGATTGTGATCGATTGGCATTTGCCCATATTCCAGCAGTGG 909
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
DB 910 TGTGGGATTAACTCATCTTTAACTACGCGCAGAGATTTCCTCGCGCGGG----- 963
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
DB 964 -----ITCGATATTACAGCAGCAGCTGAATTCGATCGCGACGGCGTC 1008
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
DB 1009 GTTAATCTGTCTTTACCATTGCG-----GCGCTGCGCTGTGGTAAATCCGGT 1059
QY 654 ArgArg---PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
DB 1060 CGCGGTAAATTAATGCTGTTCGGCTTCGGGATTG-----ACGCTGATCTATGTG 1110
QY 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
DB 1111 CTGATCGCGCGGCTTACGCCATGGCATTTATGGGTGGCGCGTACTG-----TTGCTG 1164
QY 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
DB 1165 GTCTGCGCGCGATTGCTATTATGCTTACCCCTGCGCGCGGTGACCTGGGTGCTGCTG 1224
QY 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
DB 1225 GCGGAGATTTTCCCAACCGGTACGTGGTACTCGCATGCTTTTAGGTACCTTGGCGCTG 1284
QY 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
DB 1285 TGATCGCTGTTTCCTGTTAACTATACCTTCCCGCTGCTTAATGCCGCTCGGGCGCG 1344
QY 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
DB 1345 GCAGGAAGCTTCTGCTGTATGGTGTATTATGGCTGCCGGCTACCTCTATATCTCGCGC 1404
QY 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAla 790
DB 1405 AACGTCCCGAAGCGAAGGTATACCCCTCGAAGCGCTGGAAGACGAGCTGGCG 1458
```

## RESULT 15

US-09-489-039A-2378  
; Sequence 2378, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 2378  
; LENGTH: 1566  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-2378

## Alignment Scores:

Pred. No.:	1,11e-34	Length:	1566
Score:	426.00	Matches:	163
Percent Similarity:	34.24%	Conservative:	88
Best Local Similarity:	22.24%	Mismatches:	176
Query Match:	10.45%	Indels:	306
DB:	4	Gaps:	20

US-10-051-909-32 (1-800) x US-09-489-039A-2378 (1-1566)

```
QY 62 ValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAla 81
DB 172 ATCACCTTAGTGCCACCTTCGCGCGCTTTCGGTTTACGACACCGCGCTTATCAAC 231
QY 82 AlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGly 101
DB 232 GCGCGCTTTCTCTACGAAGCAGTACATGGCGCTG-----ACCCCAACACCGAAGGG 285
QY 102 LeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeu 121
DB 286 CTGGTAATGAGCTTCTGCTGGCGCGCGCTGGCAGCGTGTTCGCGCAATTT 345
QY 122 SerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSer 141
DB 346 GCGGATTACTCGCGCGGTAAATATTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGGC 405
QY 142 GlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAsp 161
DB 406 GCGTTCCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 465
QY 162 GlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPro 181
DB 466 GGCTATGCGCTCGCGCGCGCTTCGGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 525
QY 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPhe 201
DB 526 ACCGAGATCGCGCGCAAGCTGACCGCGCTGAATGAAGTCGCCATCGCTCATTTGCCAGCTG 585
QY 202 LeuSerTyrCysMet-----ValPheGlyMetSerSerLeuSerProSerProAsp---Trp 218
DB 586 GCGCGCTTTGCCATCAACCGCATATTGGCATCATCTGGGGCCCATCTGCGGAGCGTCTGG 645
QY 219 ArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePhe 238
DB 646 CGGTATATCTGCTGTGTACAGCGATTTCGGCCATCTGCGCTGTTTC---GTTGGCATGTGG 702
QY 239 TyrLeuProGlySerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLys 258
DB 703 CGCGCGCGCGAAGCCCGCGCTGGCTGATCAGCAAAATCGTCATGATGAAGCGCTGCAT 762
QY 259 ValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGlu 278
DB 763 ATCTCAAAACAAATCCG-----
QY 279 GlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGlu 298
DB 781 -----CCGGCGGAACGC 792
```

```

299 AlaAlaAspLeuValThrAspGlyAspGlyGlnIleThrLeuTyrGlyProGlu 318
   |||
793 GCG-----CAGAGGAG----- 804
Qy
319 GluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeu 338
   |||
804 ----- 804
Qy
339 SerLeuAlaSerArgHisGlySerMetValaGlnSerValProLeuMetAspProIle 358
   |||
805 -----TACGATGCATC 816
Qy
359 ValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSer 378
   |||
817 TCTACCTGATC----- 828
Qy
379 ThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGlu 398
   |||
829 -----AAATTGAA 837
Qy
399 GlnTrpAspGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGly 418
   |||
838 -----GCTGGT 843
Qy
419 GlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGlu 438
   |||
844 AATAAGTAC-----ACGCCACAGACCTTTGCG-- 873
Qy
439 GlyLysAspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThr 458
   |||
873 ----- 873
Qy
459 LeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeu 478
   |||
873 ----- 873
Qy
479 AlaTrpLysTrpSerGlnLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArg 498
   |||
873 ----- 873
Qy
499 ValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuPro 518
   |||
874 -----ACGATTGTCAAAACACCC 891
Qy
519 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGln 538
   |||
891 ----- 891
Qy
539 SerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHis 558
   |||
891 ----- 891
Qy
559 ProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPheGluProGlyValArg 578
   |||
892 -----TGG-----ATCCTC 900
Qy
579 ArgAlaLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGly 598
   |||
901 AAATTGTGCTGTGGCATACCTGGCGCGCACTGCAGCAAAACCACCGCGGTTAACGTC 960
Qy
599 ValLeuTyrTrpThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLys 618
   |||
961 ATTATGTTTACGGTACGGAATCTCTGAGTCCCGCGTTTC----- 1002
Qy
619 PheGlyLeuSerSerAlaSerIleLeuIleSerSerIleThrThrLeuLeuMet 638
   |||
1003 -----TCAGAGCGTACTTCGTGATCTGCAACGTGCTGAACGGTGTTCCTCC 1050
Qy
639 Leu-----ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArg--- 655
   |||
1051 GTGGCGGTATGCTGATCGGG---GTACTGTCTCTTCGACCGCTTCAACAGTAAGACA 1107
Qy
656 -----PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670

```

```

1108 ATCATCATCTACGGTTTCGCCATTATGGCAACCTGTCATCTGATTATCGCGCTGTGCAT 1167
   |||
671 IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
   |||
1168 TACACCTGGTGGGT-----GACCTGAAAGCTACCGCCATCTGGCTGCTGGCGCGC 1218
Qy
691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
   |||
1219 CRGTTTCGTCGGCGTGATGCAGGGCTCC-----ATGGGCTTT-----ATTACCTGGGTG 1266
Qy
711 LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe 730
   |||
1267 GTATTAGCCGAACCTCTTCGCTGAAATTCGCGGACTGTGCGATGGGGATCTCGGTATTT 1326
Qy
731 ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750
   |||
1327 TTCATGTGGATCATCAACGCGGTGTCAGCTATCTGTTCCGCTACTGCAGGCAAACTG 1386
Qy
751 GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheVal 770
   |||
1387 GGCCTTGGGCCAGTCTCTTTATCTTCGCGCCCACTTAACCTATTAGCCATTCTTTTCGTG 1446
Qy
771 PheLeuLysValProGluThrLysGlyMetProLeuGlu 783
   |||
1447 GTCTTTGCCCTGCCGAGACCTCCAATAAATCGCTTGAG 1485

```

Search completed: April 13, 2005, 22:00:18  
Job time : 376.838 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 18:06:15 ; Search time 1016.41 Seconds  
(without alignments)  
4775.020 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSWLAVQTPTPLDRR.....PLEVITEFFAVGAKQAQAAKA 800

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US10051909/runat\_13042005\_074039\_14141/app\_query.fasta\_1.1678  
-DB=Published Applications NA -QFMT=fasta -SUFFIX=rnnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10051909 @CGN 1 1 953 @runat\_13042005\_074039\_14141  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	4075	100.0	2777	13	US-10-051-909-31	Sequence 31	
2	4059	99.6	2908	17	US-10-310-154-350	Sequence 350	
3	3517	86.3	2824	13	US-10-051-902-1	Sequence 1	
4	3517	86.3	2824	13	US-10-051-909-1	Sequence 1	
5	2869	70.4	2031	17	US-10-425-114-32961	Sequence 32961	
c	2835.5	69.6	2769	18	US-10-437-963-93472	Sequence 93472	
7	2788.5	68.4	4330	18	US-10-425-115-70742	Sequence 70742	
8	2674	65.6	2601	13	US-10-051-902-7	Sequence 7	
9	2674	65.6	2601	13	US-10-051-909-7	Sequence 7	
10	2584	63.4	2190	9	US-09-938-842A-1315	Sequence 1315	
11	2584	63.4	2190	11	US-09-938-842A-1315	Sequence 1315	
12	2571	63.1	2671	17	US-10-424-599-10324	Sequence 10324	
13	2543	62.9	2405	17	US-10-425-114-7958	Sequence 7958	
14	2483.5	60.4	2573	17	US-10-424-599-108435	Sequence 108435	
15	2442.5	59.9	2271	17	US-10-425-114-29146	Sequence 29146	
16	2275	55.8	2205	9	US-09-938-842A-2254	Sequence 2254	
17	2275	55.8	2205	11	US-09-938-842A-2254	Sequence 2254	
c	2245.5	55.1	2432	18	US-10-437-963-47424	Sequence 47424	
19	1963	48.2	1959	18	US-10-437-963-92959	Sequence 92959	
c	20	1763	43.3	2231	18	US-10-437-963-70909	Sequence 70909
21	1670.5	41.0	1692	13	US-10-051-902-9	Sequence 9	
22	1670.5	41.0	1692	13	US-10-051-909-9	Sequence 9	
23	1545	37.9	3205	17	US-10-424-599-77967	Sequence 77967	
24	1424	34.9	1487	13	US-10-051-902-13	Sequence 13	
25	1424	34.9	1487	13	US-10-051-909-13	Sequence 13	
26	1392	34.2	1313	18	US-10-767-701-12595	Sequence 12595	
27	1252.5	30.7	1289	17	US-10-424-599-42000	Sequence 42000	
28	1035	25.4	1412	17	US-10-260-238-747	Sequence 747	
29	959	23.5	1200	18	US-10-425-115-159851	Sequence 159851	
c	949.5	23.3	1259	17	US-10-424-599-101136	Sequence 101136	
31	943	23.1	1009	13	US-10-051-902-15	Sequence 15	
32	943	23.1	1009	13	US-10-051-909-15	Sequence 15	
33	813	20.0	761	17	US-10-425-114-35910	Sequence 35910	
34	793	19.5	650	17	US-10-425-114-3864	Sequence 3864	
35	745.5	18.3	832	18	US-10-425-115-135446	Sequence 135446	
36	663	16.3	435	18	US-10-767-701-18822	Sequence 18822	
37	620	15.2	711	18	US-10-767-701-12614	Sequence 12614	
38	620	15.2	751	17	US-10-260-238-2683	Sequence 2683	
39	615.5	15.1	2253	18	US-10-437-963-49329	Sequence 49329	
40	604	14.8	674	18	US-10-437-963-60088	Sequence 60088	
41	600	14.7	778	17	US-10-260-238-2684	Sequence 2684	
42	587.5	14.4	1374	17	US-10-369-493-47011	Sequence 47011	
43	587.5	14.4	1518	19	US-10-332-815A-10	Sequence 10	
44	580	14.2	583	18	US-10-021-323-11734	Sequence 11734	
45	562	13.8	870	13	US-10-051-902-5	Sequence 5	

#### ALIGNMENTS

RESULT 1  
US-10-051-909-31  
; Sequence 31, Application US/10051909  
; Publication No. US20020199217A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Helentjaris, Tim  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: B1163 US CIP  
; CURRENT APPLICATION NUMBER: US/10/051,909  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 60/083,044  
; PRIOR FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 31  
; LENGTH: 2777

TYPE: DNA  
ORGANISM: Zea mays  
US-10-051-909-31

Alignment Scores:  
Pred. No.: 0 Length: 2777  
Score: 4075.00 Matches: 800  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-051-909-32 (1-800) x US-10-051-909-31 (1-2777)

QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArg 20  
DB 12 ATTCGGAGCGGCTCTTGGCTTGCAGTCCAAAGCGCCCTTCAACCCCTGATCTGGACGGAGG 71  
QY 21 GluArgLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40  
DB 72 GAGCGGCTCTTCCGTCAGTGTGTTCTTGGCTGGGCGCTCTTCCGCGCTGCTTCGTGT 131  
QY 41 SerSerGlnGluProValThrSerAspAspLeuLeuGluAspLysMetSerGlyAlaVal 60  
DB 132 TCTTCACAGAGCGGTGACCTCGACGATATCTTGGAGACAAGATGTCGGGGGCTGTT 191  
QY 61 LeuValAlaLeuValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80  
DB 192 CTTGTGCGCCATAGTCGCTCCATCGGCAATCTATTGCGGGGTGGACANTGCCACCATC 251  
QY 81 AlaAlaAlaValLeuTrpIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100  
DB 252 GCAGCTGCTCTTGTATATAAAGAAAGAAATTCATTTGCAAAATGAGCCCACTGTGGAG 311  
QY 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120  
DB 312 GGACTAATTGTCAATGTCACTTATCGGGCCACCATCGTTACTACATTCCTCGGGGCA 371  
QY 121 LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTrpPhe 140  
DB 372 TTATCAGACTCGATTGGCCGACGCCCTATGCTATTATCTCTTCAATTCGTACTTCTTC 431  
QY 141 SerGlyLeuIleMetLeuTrpSerProAsnValTrpValLeuLeuAlaArgPheVal 160  
DB 432 AGCGGCTCATATGCTATGCTCTCTTATGCTATGCTGCTGTTGGCAGCTTCGTA 491  
QY 161 AspGlyPheGlyIleGlyAlaValThrLeuValProLeuTrpIleSerGluIleAla 180  
DB 492 GATGGATTGTGATTTGGCTTGGCTGTCCACCTTGTCCCTTTGTACATTTCAAGAAATAGCC 551  
QY 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200  
DB 552 CCTTCGGAGATTAGAGTTTGTCTGAATACACTACCAATTCAGTGGATCAGAGGAATG 611  
QY 201 PheLeuSerTrpCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220  
DB 612 TTCTTGTCACTATGATGATGTTGGATGCTCCCTGTGCGCAATCACCCGATTGGAGAAT 671  
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTrpLeu 240  
DB 672 ATGCTTGGTGTCTCGGATACCTTCAATGTTCTTCTTGGTTTGAATAATTTATCTCT 731  
QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260  
DB 732 CCTGAATCTCCAGATGGCTGTAGCAAGGTCCGATGCGACAGCAAAAAGGTGTGTG 791  
QY 261 GlnLysLeuArgGlyLysAspAspValSerGlyLeuSerLeuLeuLeuGluGlyLeu 280  
DB 792 CAAAAGTTTACGGGGGAAGACGATGCTCAGGTGAATGTCCTTCTTCGAGGGGTG 851  
QY 281 GluValGlyGlyAspThrSerIleGluGluTrpIleIleGlyProAlaThrGluAlaAla 300  
DB 852 GAGGTGGAGGAGACACTTCATTTGAAGAGTACATCAATGGACCTGCGCCCGGCGAGCC 911

QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTrpGlyProGluGluGly 320  
DB 912 GATGATCTTGTGTACTGACGGTGATAAGGAACAAATCACACTTTATGGCCCTGAAGAAGGC 971  
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340  
DB 972 CAGTCATGATGTCTCGACCTTCTAAGGAGCCCATCATGCTTGAAGGTGTCTTCTCTT 1031  
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360  
DB 1032 GCATCTCGTCATGGGAGCATGTTGAACACAGAGTGTACCCCTTATGGATCCGATGTGACA 1091  
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380  
DB 1092 CTTTTTGTGTGTCCATGAGATATGCTCAAGCTGGAGGAAGTATGAGGAGCACATTG 1151  
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400  
DB 1152 TTTCCAAACTTTTGGAGATATGTTCAAGTGTCAAGATCAGCATGCCAAAATGAGCAGTGG 1211  
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTrpAlaSerAspGlyAlaGlyAsp 420  
DB 1212 GATGAAGAGAATCTTCATAGGATGACGAGGAGTACGCATCTGATGGTGCAGAGGTGC 1271  
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440  
DB 1272 TATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGCGAGGCAACAGGTGCGGAAGGGAAG 1331  
QY 441 AspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460  
DB 1332 GACATTTGTCCACCATGGTCACCGTGAAGTGTCTTGAAGCATGAGAAGGCAAAACCTCTTA 1391  
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480  
DB 1392 GGGGAGGTTGAGATGTTGTGACGACACTGATATCGTGGGGATGGCAGCTTCTTGG 1451  
QY 481 LysTrpSerGlnLysGlyGluAsnGlyAspLysGlyGlyPheLysArgValTrp 500  
DB 1452 AAATGGTTCAGAGAGAGAGAGTGTAGTGTAGAAAGAGGTGGTTTCAAAGAGTCTAC 1511  
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520  
DB 1512 TTGCACCAAGAGGAGTCTCTGGCTCAAGAGGGCTCAATTTGTTTCACTTCCCGGTGT 1571  
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540  
DB 1572 GGCGATGTTTTTGAAGGTGTAGTGTGTACATGCTGCTGTCTTAGTAAGTCAGTCAGCA 1631  
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560  
DB 1632 CTTTTCTCAAGGGTCTTTGCTGAACCAACGATGTCAGATGCTGCCATGGTTTCAACCATCT 1691  
QY 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAla 580  
DB 1692 GAGTAGTGCCTCAAGGTTCCACGTTGGAAAGATTTGTTGAACCTGGAGTGGCGGTGCC 1751  
QY 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeu 600  
DB 1752 CTGTTAGTCCGTTTGGAAATTCAGATCTTCAACAGTTGCTTGGAAATAAACCGTGTCTG 1811  
QY 601 TyrTrpThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPheGly 620  
DB 1812 TACTATACCCCAAAATCTTGGACAGCTGGTGTGGCAGTTATCTTTCCAAATTTGTT 1871  
QY 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuMetLeuPro 640  
DB 1872 CTCAGCTCGGCATCAGCATCATCTTGTATCAGTTCTCTCACTACCTTACTTAATGCTTCT 1931  
QY 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660  
DB 1932 TGCATTTGGCTTTTGGCATGCTGCTTATGGATCTTTTCCGGAAGAAGGTTTTTGTGCTAGGC 1991

```
QY 661 ThrIleProIleuIleAlaSerLeuValIleuValValSerAenLeuIleAapLeu 680
Db 1992 ACAATTCCAATCTGTAGATCTCTAGTTATCTCTGGTTGTGCCAATCTAAATGATTG 2051
QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Db 2052 GGTACACTAGCCCATGTTTGGCTCTCCACCGTCAGTGTATGCTACTTCTGCTGCTTC 2111
QY 701 ValMetGlyPheGlyProIleProAenIleLeuCysAlaGluIlePheProThrArgVal 720
Db 2112 GTTATGGAGTTTGGTCCCATCCCAACATTTATGTGCAGAGATCTTTCCAACAGGGTT 2171
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheThrIleGlyAapIleIleValThr 740
Db 2172 CGTGGGCTCTGTATGTCATTTGTGCCTTTTACATTTCTGGATCGGAGATATCATCGTCACC 2231
QY 741 TyrSerLeuProValMetLeuAenAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db 2232 TACAGCTTCTGTGATGCTGAATGCTATTTGGACTGGCGGGTGTTCAGCATATATGCA 2291
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 2292 GTCGTATGCTTGATTTCTTTGTTGCTCTCTTAAGGTCCCTGAGACAAAGGGGATG 2351
QY 781 ProLeuGluValIleThrGluPheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 2352 CCCCTTGAGGTTATTACCGAATCTTTTGCAGTTTGGTGCAGCAAGCGGCTGCAAAAGCC 2411

RESULT 2
US-10-310-154-350
; Sequence 350, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhao
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihahieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K. R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanguo
; APPLICANT: Xu, Nanfei
```

```
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15 (52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 350
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (331)..(2565)
; OTHER INFORMATION:
; US-10-310-154-350

Alignment Scores:
Pred. No.: 0 Length: 2908
Score: 4059.00 Matches: 797
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 17 Gaps: 0

US-10-051-909-32 (1-800) x US-10-310-154-350 (1-2908)
QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAapLeuAapArgArg 20
Db 166 ATTCGGAGCGGCTCTTGGCTTGCAGTCCAGACGCCCTTACCCCTCGATCTGGACCGGAGG 225
QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
Db 226 GAGCGGCTCTTCCGTCAGTTGTTCTTGGCTTGGGCTCTTCGCGCTGCTTCGTGT 285
QY 41 SerSerGlnGluProValThrSerAapIleLeuGluAapLysMetSerGlyAlaVal 60
Db 286 TCTTCACAGAGCGGCTGACCTCGGACGATATCTTGGAGACCAAGATGTCGGGGCTGTT 345
QY 61 LeuValAlaIleValAlaSerIleGlyAenLeuLeuGlnGlyTrpAapAenAlaThrIle 80
Db 346 CTGTGCGCATAGTCGCCCTCCATCGCAATCTATTGAGGGGTGGGACAAATGCCACCATC 405
QY 81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAenGluProThrValGlu 100
Db 406 GCAGCTGCTGTTCTGTATATAAAGAGGAATTTCAATTGCAAAATGAGCCCTGTGGAG 465
QY 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120
Db 466 GGACTAATGTGTCAATGTCATTATCGGGCCACCATCGTTACTACTTCTCGGGCCA 525
QY 121 LeuSerAapSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
Db 526 TTATCAGACTCGATTGCGCGAGCGCCCTATGCTTATTTCTCTCAATTCTGTACTTCTTC 585
QY 141 SerGlyLeuIleMetLeuTrpSerProAenValTyrValLeuLeuAlaAapPheVal 160
Db 586 AGCGGCTCATCATGCTATGGTCTCCTAATGTCTATGTCTGTGTGGCAGCCTTCGTA 645
QY 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
Db 646 GATGGATTTGGTATTGGCTTGGCTGTACGCTTGTGCTTGTACATTTTCAGAAATAGCC 705
QY 181 ProSerGluIleArgGlyLeuLeuAenThrLeuProGlnPheSerGlySerGlyMet 200
Db 706 CCTTCGAGATTAGAGGTTTGTGATACATACATACATTCAGTGGATCAGGAGGAATG 765
QY 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAapTrpArgIle 220
```

Db 766 TTTCTGTCTACTGATGATGTTGGAGTCCCTGTCCTCCATCCACCGATTGGAGATT 825  
Qy 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240  
Db 826 ATGCTTGGTGTCTCGCGATACCTTCATATGTTCTCTTTGGTTTGACAAATATTTATCTT 885  
Qy 241 ProGluSerProArgTyrLeuValSerIysGlyArgMetAlaGluAlaLysIleValLeu 260  
Db 886 CCTGAATCTCCAAAGATGGCTGTTAGCAAGATCGGATGGCAGAGGCAAAAAGGTGTG 945  
Qy 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280  
Db 946 CAAAAGTTACGGGGGAAAGACCATGCTCAGGTGAATTTGCTCTTCTCTCGAAGGGTGTG 1005  
Qy 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300  
Db 1006 GAGGTTGGAGGAGACATCTCCATTGAAGAGTATACATCTTGGACCTGCCACCGAGGCGAGCC 1065  
Qy 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320  
Db 1066 GATGATCTTGTACTGACCGGTGANTAGGAACCAATCACATTTATGGGCGCTGAAGAGGC 1125  
Qy 321 GlnSerTrpIleAlaArgProSerIysGlyProIleMetLeuGlySerValLeuSerLeu 340  
Db 1126 CAGTCATGATGCTCGACCTTCCAAAGGACCCAGCATGCTTGGAGTGTGCTTCTCTT 1185  
Qy 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360  
Db 1186 GCATCTCGTCATGGGAGCATGGTGAACCAAGAGTGTACCCCTTATGGATCCGANTGTGACA 1245  
Qy 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380  
Db 1246 CTTTTTGGTACTCTCCATGAGATATGCTCTCAAGCTGGGAGGAAGTATGAGGAGCACATG 1305  
Qy 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400  
Db 1306 TTTTCCAACTTTGGAAGTATGTTCACTGTACACAGATCAGCATGCCAATAAATGAGCAGTGG 1365  
Qy 401 AspGluLeuAsnLeuHisArgAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp 420  
Db 1366 GATGAGAGAGATCTTCATAGGATGACAGGAGTACGCATCTGATGGTGCAGAGGTGAC 1425  
Qy 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440  
Db 1426 TATGAGGACAACTCTCCATAGCCCTTCTGCTCCAGGCAGCAACAGGTGCGAGAGGAG 1485  
Qy 441 AspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460  
Db 1486 GACATGTGCACCATGCTCACCGTGAAGTGTCTTGAGCATGAGAGGCAAGCCCTCTTA 1545  
Qy 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480  
Db 1546 GGGGAGGTTGAGATGTTGTGAGCAGCAGCTGATATCGGTGGGGATGGCAGCTTGTCTGG 1605  
Qy 481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500  
Db 1606 AAATGGTCAGAGAAGGAAGTGAAGATGGTAGAAGAAAGGTGGTTTCAAAAGAGCTTAC 1665  
Qy 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520  
Db 1666 TTGCACCAAGAGGAGTTCCTGGCTCAAGAGGGGCTCAATTTGTTCACTTCCCGTGTGT 1725  
Qy 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540  
Db 1726 GCGCATGTTCTTGAAGGTAGTGAAGTTTGTACATGCTGCTGCTTGTAGTAGTCAGTCAGCA 1785  
Qy 541 LeuPheSerIysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560  
Db 1786 CTTTTTCAAAGGTCTTGTGTAACCAACCGCATGTGCATGCTGCCATGTTCCCATCT 1845  
Qy 561 GluValAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAla 580

Db 1846 GAGGTAGCTGCCAAAGTTACGTTGGAAAGATTGTTTGAACCTGGAGTGAGCGGTGCC 1905  
Qy 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600  
Db 1906 CTGTTAGTCGGTGTGGAAATTCAGATCCTTCAACAGATTGCTGGAATAAACCAGGTGTTCTG 1965  
Qy 601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620  
Db 1966 TACTATACCCCAAAATCTTGGCAAGCTGCTGTGGCAGTTATTTCTTCCAAATTTGGT 2025  
Qy 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640  
Db 2026 CTGAGCTCGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCT 2085  
Qy 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgAspPheLeuLeuGly 660  
Db 2086 TGCAATGGCTTTGCCATGCTTATGGATCTTCCGAAGAGGTTTTTGTCTGCTAGGC 2145  
Qy 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680  
Db 2146 ACAATTCCAATCTTGATAGCATCTAGTTATCTGGTGTGTGTCATCTAATGATTTG 2205  
Qy 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValIleValIlePheCysPhe 700  
Db 2206 GGTACACTAGCCCATGCTTGTCTCTCCACCGTCAGTGTATGCTCTACTTCTGCTGCTTC 2265  
Qy 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720  
Db 2266 GTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCAGAGATCTTTTCCAAACCAAGGTT 2325  
Qy 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740  
Db 2326 CGTGGGCTCTGTATTTGCCATTTTGCCCTTACATTTGATCGAGAGATATCATGTCACC 2385  
Qy 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTrpAla 760  
Db 2386 TACAGCTTCTCTGTGATGCTGAATGCTATTTGACTGGCGGGTGTTCAGCATATATGCA 2445  
Qy 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780  
Db 2446 GTCGTATGCTTGATTTCTTGTGTTGCTTCTTAAAGTCCCTGAGCAAGGGGATG 2505  
Qy 781 ProLeuGluValIleThrGluPhePheAlaValAlaGlyAlaLysGlnAlaAlaLysAla 800  
Db 2506 CCCCTTGAGGTTATTATCCGAATTTCTTTGAGTTGGTGGAGCAAGCGGCTGCAAAAGCC 2565

RESULT 3  
US-10-051-902-1  
; Sequence 1, Application US/10051902  
; Publication No. US20020178468A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Kinney, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BS-1163  
; CURRENT APPLICATION NUMBER: US/10/051,902  
; PRIOR APPLICATION DATE: 2002-01-17  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: US/09/291,922  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 2824  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (29)  
; NAME/KEY: unsure  
; LOCATION: (622)  
; NAME/KEY: unsure



; LOCATION: (636)  
 ; NAME/KEY: unsure  
 ; LOCATION: (638)  
 ; NAME/KEY: unsure  
 ; LOCATION: (669)  
 ; NAME/KEY: unsure  
 ; LOCATION: (711)  
 ; NAME/KEY: unsure  
 ; LOCATION: (822)  
 ; NAME/KEY: unsure  
 ; LOCATION: (856)  
 ; NAME/KEY: unsure  
 ; LOCATION: (889)  
 ; NAME/KEY: unsure  
 ; LOCATION: (896)  
 ; NAME/KEY: unsure  
 ; LOCATION: (944)  
 ; NAME/KEY: unsure

US-10-051-902-1

# Alignment Scores:

Pred. No.: 0 Length: 2824  
 Score: 3517.00 Matches: 684  
 Percent Similarity: 95.18% Conservative: 27  
 Best Local Similarity: 91.57% Mismatches: 34  
 Query Match: 86.31% Indels: 2  
 DB: 13 Gaps: 2

US-10-051-909-32 (1-800) x US-10-051-902-1 (1-2824)

QY 56 MetSerGlyValAlaValLeuValAlaValAlaSerIleGlyAenLeuGlnGlyTyr 75  
 DB 238 ATGGGGGCGCGGTGATGGTGGCCATCGCGCCCTCTATCGCACTTGTGCGAGGGCTGG 297  
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
 DB 298 GACATGCGACATTTGCTGAGCGCTCTGTACATATAAGAGGAATTCACCTTCGACAGC 357  
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
 DB 358 GAGCCTCTGATCGAAGGCCCTCATGCGCCATGTTCTCATTTGGGGCAACAGTCATCACA 417  
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
 DB 418 ACATCTCCGGGGCCAGGGCTGATGCTGGTGGTAGGAGGCCCATGCTGGCTCGCTCGCT 477  
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155  
 DB 478 GTCTCTACTTGTGTCAGTGGGCTGGTATGCTTTGGGCGCAATTTGTTACATCTTGCTC 537  
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
 DB 538 CTCGAAGGCTCATTTGATGGTTCGGTATCGGTTTGGCGGTTCACACTTGTTCCTCTCTAC 597  
 QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194  
 DB 598 ATCTCCGAAATCGCACCGCACAGANATTTCTTGGGGCTGTNGAACACGTTGGCGCAGTTC 657  
 QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213  
 DB 658 ATTGGGGTCAGNGAGGAGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 717  
 QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe 233  
 DB 718 CCCAAACCTGATTGGAGGCTCATGTTGGAGTTCGTGCGATCCGCTCACCTTATNTACTTT 777  
 QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253  
 DB 778 GGAATGACTGCTCTCTACTTTCCTCTGATATCCAGAGTGGCTTGTNAGCAAGGAGGATG 837  
 QY 254 AlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273  
 DB 838 GCGAGGCGAAGAGAGTGTGCAAGGCTCGGGGAAGAGAGATGTCTCANGGGGAGANG 897

QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293  
 DB 898 GCTCTTCTAGTTGAAGGTTTGGGGGTCGGTAAAGATACACGTATTTNAGAGTACATCAT 957  
 QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313  
 DB 958 GGACCTGCCACCGAGGCGCGGATGATCTTGTAACTGACGGTGTAAAGAAACAATACACA 1017  
 QY 314 LeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333  
 DB 1018 CTTTATGGGCTCGAAGAGGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG 1077  
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353  
 DB 1078 CTTTGAAGTGTGCTTTCTCTTCATCTCGTCATGGAGCATGTGTAAACACAGAGTGTACCC 1137  
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAenMetProGlnAlaGly 373  
 DB 1138 CTTATGGATCCGATTTGTGACACTTTTGTGTAGTGTCTCATGAGANTATGCCTCAAGCTGGA 1197  
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
 DB 1198 GGAAGTATGAGGAGCACATTTGTTTCCAAACTTTTGAAGTATGTTCAGTGTACAGATCAG 1257  
 QY 394 HisAlaLysAsnGlnGlnTrpAspGluGluAenLeuHisArgAspAspGluGluTyrAla 413  
 DB 1258 CATGCCAAAATGAGCAGTGGGATGAAGAGAATTTTCATAGGATGACGAGGAGTACGCA 1317  
 QY 414 SerAspGlyValaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerAspGln 433  
 DB 1318 TCTGATGTTGAGGAGGTGACTATAGGACATCTCATAGCCCATTTGCTGTCCAGGCAG 1377  
 QY 434 AlaThrGlyValaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453  
 DB 1378 GCAACAGGTGCGAAGGAAGGACATTTGTGCACCATGTCACCGTGAAGTGTCTTGAGC 1437  
 QY 454 MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGly 473  
 DB 1438 ATGAGAAGCCAAAGCTCTTTAGGGGAGGTGGAGATGGTGTGAGCAGCACTGTATCCGT 1497  
 QY 474 GlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAenGlyValGlyGlu 493  
 DB 1498 GGGGATGCCAGCTTCTTGGAAATGTCAGAGAGGAGGTGAGATGTTGTAGNAAGGAA 1557  
 QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513  
 DB 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCTCTGGCTCAACAAGGGCTCA 1617  
 QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533  
 DB 1618 ATTTGTTTCACTTCCCGTGTGGCGATGTTCTTTGAGGGTAGTGAGTTTGTACATGCTGCT 1677  
 QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
 DB 1678 GCTTTTATAGTCAGTCAGCACCTTTTCTCAAGGGTCTTCTGTAACCCACCATGTCAGAT 1737  
 QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573  
 DB 1738 GCTGCCATGGTTCAACCCATCTGAGGTAGCTGCCAAAGGTTTCAGTTGGAAAGATTTGTTT 1797  
 QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593  
 DB 1798 GAACCTGGAGTGGAGGGTGCCCTGTTAGTCGGTGTGGAAATTCAGATCTCTCAACAGTTT 1857  
 QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613  
 DB 1858 GCTGGAAATAAACCGGTGTTCTGTACTATATCCCAAAATTTCTTGAGCAAGCTGGTGTGGCA 1917  
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu 633  
 DB 1918 GTTATTTCTTCCAAATTTGGTCTCAGCTCGGCATCAGATCCATCTTGATCAGTTCTCTC 1977  
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653

Db 1978 ACTACCTTAATAAGTCTCTTGGCATGCTGCTTGGCATGCTTATGGATCTTTCGGGA 2037  
Qy 654 AtgAqgPheLeuLeuLeuGlyThrIleProIleLeuLeuAlaSerLeuValIleLeuVal 673  
Db 2038 AGAAGGTTTTTCTGCTAGGACAAATTCATCTTTGATAGCATCTCTAGTTATCTCGTT 2097  
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
Db 2098 GTGTCCAACTAATGATTTGGGTACATAGCCCATGCTTGTCTCCACCATCAGTGT 2157  
Qy 694 IleValThrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713  
Db 2158 ATCGTCTACTTCTGCTGCTTCTGTTATGGGATTTGTCCTCCATCCCAACATTTAAGTCA 2217  
Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733  
Db 2218 GAGATCTTTCCACACCAAGGGTTCGTGGCTCTGTATTTGCCATTTGTGCTTTACATTCGG 2277  
Qy 734 IleGlyAspIleIleValThrTrpSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753  
Db 2278 ATCGGAGATATCATCTCACCTACAGCTTCTCTGTGATGCTGAATGCTATTGGAGCTGGCG 2337  
Qy 754 GlyValPheSerIleTrpAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
Db 2338 GGTGTTTTTCAGCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2397  
Qy 774 ValProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793  
Db 2398 GTCCCTTGACAAAGGGGATGCCCCCTTGAGTTATACCGAATCTTTTGCGATTTGGTGGC 2457  
Qy 794 LysGlnAlaAlaLysAla 800  
Db 2458 AGCAAGCGGCTGCAAAAGCC 2478

RESULT 4

US-10-051-909-1  
Sequence 1, Application US/10051909  
Publication No. US20020199217A1  
GENERAL INFORMATION:  
APPLICANT: Allen, Steve  
APPLICANT: Helentjaris, Tim  
APPLICANT: Hitz, Bill  
APPLICANT: Kinney, Tony  
APPLICANT: Tingey, Scott  
TITLE OF INVENTION: Plant Sugar Transport Proteins  
FILE REFERENCE: BRL163 US CIP  
CURRENT APPLICATION NUMBER: US/10/051,909  
CURRENT FILING DATE: 2002-01-17  
PRIOR APPLICATION NUMBER: 60/083,044  
PRIOR FILING DATE: April 24, 1998  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 1  
LENGTH: 2824  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (29)  
NAME/KEY: unsure  
LOCATION: (622)  
NAME/KEY: unsure  
LOCATION: (636)  
NAME/KEY: unsure  
LOCATION: (638)  
NAME/KEY: unsure  
LOCATION: (659)  
NAME/KEY: unsure  
LOCATION: (771)  
NAME/KEY: unsure  
LOCATION: (822)  
NAME/KEY: unsure

; LOCATION: (856)  
; NAME/KEY: unsure  
; LOCATION: (889)  
; NAME/KEY: unsure  
; LOCATION: (896)  
; NAME/KEY: unsure  
; LOCATION: (944)  
US-10-051-909-1  
Alignment Scores:  
Pred. No.: 0 Length: 2824  
Score: 3517.00 Matches: 684  
Percent Similarity: 95.18% Conservative: 27  
Best Local Similarity: 91.57% Mismatches: 34  
Query Match: 86.31% Indels: 2  
DB: 13 Gaps: 2  
US-10-051-909-32 (1-800) x US-10-051-909-1 (1-2824)  
Qy 56 MetSerGlyAlaValLeuValIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
Db 238 ATGGGGGGGGCGCGTGTGTCGCCATCGGGCTCTATCGCACTTGTCTGAGGGCTGG 297  
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTrpIleLysLysGluPheGlnLeuGlnAsn 95  
Db 298 GACAATGCCACAATTCGTGGAGCCGTCGTGTACATAAAGAAGGAATTCACCTGCAGAG 357  
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Db 358 GAGCCTCTGATCGAAGGCCCTCATCTGTCGCCATGTTCTCATTTGGGGCAACAGTCATCA 417  
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
Db 418 ACATCTCCGGGGCCAAAGGCTGACTGCTGTGTAGAGGCCCATGCTGCTGCGCTCGGCT 477  
Qy 136 IleLeuTrpPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTrpValLeuLeu 155  
Db 478 GTCCCTTACTTGTCTAGTGGGCTGGTGTGTCCTTGGGGCCCAATTTGTGTACATCTTGTCT 537  
Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTrp 175  
Db 538 CTCGAAGGCTCATGATGGGTTTCGTTATCGGTTTGGCGGTCACACTTGTCTCTCTAC 597  
Qy 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrIleProGlnPhe 194  
Db 598 ATCTCGAAACTGCACCCGACAGANATCTTGGGGCTGNTNGAACACCGTTGCCGCGATTC 657  
Qy 195 SerGly---SerGlyGlyMetPheLeuSerTrpCysMetValPheGlyMetSerLeuSer 213  
Db 658 ATTGGGGTTCAGNGGAGGAGTGTCTCTCTACTGCTGCTGTTTGGGATGTCCTCATG 717  
Qy 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe 233  
Db 718 CCCAAACCTGATTGGAGGCTCATGCTTGGAGTTCTGTGATCCCGCTCACTTATNTACTTT 777  
Qy 234 GlyLeuThrIlePheTrpLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253  
Db 778 GGACTGACTGTCTTCTTCTGCTGAATCCCAAGGCTGCTTGTNAGCAAGAGGATG 837  
Qy 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeu 273  
Db 838 GCGAGCGCGAAGAGAGTGTGCAAGGCTCGGGGAGAGAAGATCTCTCANGGGAGANG 897  
Qy 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTrpIleIle 293  
Db 898 GCTCTTCTAGTTGAAGGTTTGGGGGTCGGTAAAGATACACGATTTTNAGAGTACATCAT 957  
Qy 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313  
Db 958 GGACCTGCCACCGAGCGAGCCGATGATCTTGTAACTGACGGTGATAGAGAAACAAATCA 1017  
Qy 314 LeuTrpGlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProfileMet 333

Db 1018 CTTTATGGCCCTGAAGAGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG 1077  
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAenGlnSerValPro 353  
Db 1078 CTTGGAAGTGTGCTTTCTCTTGCACTCTCGTCATGGGAGCATGTTGAACAGAGTGTACCC 1137  
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAenMetProGlnAlaGly 373  
Db 1138 CTTTATGGATCCGATTTGTACACTTTTGGTAGTGTCCATGAGAATATGCCCTCAAGCTGGA 1197  
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
Db 1198 GGAAGTATCAGGAGCACATTTGTTCCAACTTTTGGAAAGTATGTTTCAGTGTCCAGATCAG 1257  
QY 394 HisAlaLysAenGluGlnTrpAspGluGluAenLeuHisArgAspAspGluGluTrpAla 413  
Db 1258 CATGCCAAAATGAGCAGTGGGATGAAGAGAAATCTTCATAGGATGACGAGGAGTACGCA 1317  
QY 414 SerAspGlyValGlyGlyAspTyrGluAspAenLeuHisSerProLeuLeuSerArgGln 433  
Db 1318 TCTGATGGTGCAGAGGTGACTATAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCAG 1377  
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453  
Db 1378 GCACAGGTGCCGAGGAGGAGACATTTGTGCACATGGTCACCTGGAGTGTCTTGAGC 1437  
QY 454 MetArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473  
Db 1438 ATGAGAAGGCAGAACCTCTTAGGGAGGGTGGAGATGGTGTGAGCAGCACTGTATCGGT 1497  
QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAenGlyArgLysGlu 493  
Db 1498 GGCGGATGGCAGCTTCTTGGAAATGGTTCAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAA 1557  
QY 494 GlyGlyPheLysArgValTrpLeuHisGlnGluGlyValProGlySerArgArgGlySer 513  
Db 1558 GGTGGTTTCAAAGAGTCTACTTGCACCAAGAGGGATTCCTGGCTCAAGAGGGGCTCA 1617  
QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533  
Db 1618 ATTGTTTCACTTCCCGGTGGTGGCATGTTCTTGGGGTAGTGTGAGTTGTGTACATGCTGCT 1677  
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
Db 1678 GCTTTTAGTACGTCAGCACTTTTCTCAAGAGGGTCTTGTCAACCCAGCATGTCCAGAT 1737  
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573  
Db 1738 GCTGCCATGTTCACCCATCTGAGGTAGCTGCCAAAGGTTCACTGTGGAAAGATTGTTT 1797  
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593  
Db 1798 GAACCTGGAGTAGGGTGGCTGTTAGTGGTGGTGGATTTCAGATCCTTCAACAGTTT 1857  
QY 594 AlaGlyIleAenGlyValLeuTrpThrProGlnIleLeuGluGlnAlaGlyValAla 613  
Db 1858 GCTGGAATAAACGGTGTCTGTACTATACCCCAAAATTTCTGAGCAAGTGTGTGGCA 1917  
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633  
Db 1918 GTTATTCTTTCCAAATTTGGTCTCAGCTCGGCATCAGATCCCATCTTGATCTCTCTC 1977  
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653  
Db 1978 ACTACCTTACTAATGCTTCTTGCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2037  
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
Db 2038 AGAAGGTTTTTGTGCTAGGCACAAATTCCTGATGATGATCTCTAGTATCTCTGTT 2097  
QY 674 ValSerAenLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
Db 2098 GTGTCCAATCTAATTTGTTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157

QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProIleProIleLeuCysAla 713  
Db 2158 ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGA 2217  
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733  
Db 2218 GAGATCTTTCCAAACNAGGGTTGTCGCCCTCTGTATTGCCATTTGTGCCCTTACATTCGG 2277  
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAenAlaIleGlyLeuAla 753  
Db 2278 ATCGGAGATATCATCTCACCTACAGCTTCCTGTGATGCTGAATGCTATTGGACTGGCG 2337  
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
Db 2338 GGTGTTTTTTCAGCATATATGAGTCGTATCTTGATTTCTTTGTGTCTTCCTTAAG 2397  
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793  
Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGTTATTACCGAATTTCTTGCAGTTGGTGGCG 2457  
QY 794 LysGlnAlaAlaAlaLysAla 800  
Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478  
RESULT 5  
US-10-425-114-32961  
; Sequence 32961, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 32961  
; LENGTH: 2031  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17045C09\_FLI  
US-10-425-114-32961

Alignment Scores:  
Pred. No.: 3,75e-274 Length: 2031  
Score: 2869.00 Matches: 561  
Percent Similarity: 99.65% Conservative: 1  
Best Local Similarity: 99.47% Mismatches: 2  
Query Match: 70.40% Indels: 0  
DB: 17 Gaps: 0

US-10-051-909-32 (1-800) x US-10-425-114-32961 (1-2031)

QY 237 IlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAla 256  
Db 2 ATATTTTATCTCTCTGAAATCTCCAGATGGCTGTAGCAAGGTCCGATGGCAGAGGCA 61  
QY 257 LysLeuValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeu 276  
Db 62 AAAAAGGTGTGCAAAAGTTTACGGGGAAAGACCATGTCTCAGTGAATTTGCTCTCTT 121  
QY 277 LeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTrpIleLeuGlyProAla 296  
Db 122 CTCGAAGGTTGGAGGTGGAGGAGACACTTCCATTGAGAGTACATCATTTGGACCTGCC 181  
QY 297 ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTrpGly 316

```

Db 182 ACCGAGGCGCCGATGATCTTCTTACTGCGGTGATAGGAACAAATCACACTTTATGGG 241
Qy 317 ProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySer 336
Db 242 CTGGAAGAAGCCAGTCATGATGCTGACCTTCCAGGGACCCAGCATGCTTGAAGT 301
Qy 337 ValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAsp 356
Db 302 GTGCTTTCTTGTGCATCTCGTCATGGAGCATGGTGAACACAGAGTACCCCTTATGGAT 361
Qy 357 ProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMet 376
Db 362 CCGATTGTGACACTTTTGTGTAGTGTCCATGAGATATGCTCAAGCTGGAGGAAGTATG 421
Qy 377 ArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLys 396
Db 422 AGGAGCACATTTGTCAAACTTTGGAAGTATGTTTCAGTGTACAGATACGATGCCAAA 481
Qy 397 AsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGly 416
Db 482 AATGACGACGTGGATGAAGAGAATCTTCATAGGATGACAGGAGTACGCATCTGTATGGT 541
Qy 417 AlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGly 436
Db 542 GCAGGAGGTGACTATGAGGACCAATCTCCATAGCCCAATTCCTGCTCCAGGCGAGCAACAGGT 601
Qy 437 AlaGluGlyAspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArg 456
Db 602 GCGGAAGGAGGACATGTCACCATGTCACCGTGGAGTGTCTTGGAGCATGAGGAAGG 661
Qy 457 GlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTyr 476
Db 662 CAAAGCCTCTTAGGGAGGCTGGAGATGCTGTGAGCAGCAGCTGATATCGGTGGGGAGTG 721
Qy 477 GlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPhe 496
Db 722 CAGCTTGTCTGAAATGGTCAGAGAAGGAGTGAATGTTAGATGGTAGAAGAGGTTGTTTC 781
Qy 497 LysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSer 516
Db 782 AAAAGAGTCTACTTGCCACCAAGAGGAGTTCCTGGCTCAAGAAGGGGCTCAATGTTTCA 841
Qy 517 LeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuVal 536
Db 842 CTTCCCGGTGTGGCGATGTTCTTGAGGGTGTAGTGTGTTGATGCTGCTTTAGTA 901
Qy 537 SerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMet 556
Db 902 AGTCAGTCAGCACTTTTCTCAAGGGTCTTGTCTGAACCAAGCATGTCAGATGCTGCCATG 961
Qy 557 ValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGly 576
Db 962 GTTCACCCCATCTGAGGTAGCTGCCAAGGTTCCAGCTTGAAGATTTGTTGAACCTGGA 1021
Qy 577 ValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIle 596
Db 1022 GTGAGGCGTGCCTGTGTAGTCGGTGTGGAATTCAGATCTTCAACAGTTTGTGGGAATA 1081
Qy 597 AsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeu 616
Db 1082 AACGGTGTCTGTACTATATACCCACAAATCTTGAAGAGCTGGTGGCGATTTATCTT 1141
Qy 617 SerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeu 636
Db 1142 TCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTCTCTCACTACCTTA 1201
Qy 637 LeuMetLeuProCysIleGlyPheAlaMetLeuMetAspLeuSerGlyArgArgPhe 656
Db 1202 CTATGCTTCTTGGATGGCTTGGCTTGGCATGCTGCTTATGATCTTTCCGGAAGAGGTTT 1261
Qy 657 LeuLeuLeuGlyThrIleProIleIleIleAlaSerLeuValIleLeuValValSerAsn 676

```

```

Db 1262 TTCTGCTAGGACAAATTCATCTTGTATGATGATCTCTAGTATCTCTGGTGTGTCCAAT 1321
Qy 677 LeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyr 696
Db 1322 CTAATTGATTGGGTACACATAGCCCATGCTTTGTCTCTCCACCGTCAGTGTATCGTCTAC 1381
Qy 697 PheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhe 716
Db 1382 TTCTGCTGCTTGTATGGGATTTGGTCCCATCCCAACATTTTATGTGACAGATCTTT 1441
Qy 717 ProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAsp 736
Db 1442 CCAACACAGGTTCTGGGCTCTGTATTCATTTGTGCTTTTACATCTCGATCGAGAT 1501
Qy 737 IleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPhe 756
Db 1502 ATCATCGTCACCTACAGCTTCTGTGTGATGCTGAATGCTATGTGACTGGCGGGTGTTC 1561
Qy 757 SerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLeuValProGlu 776
Db 1562 AGCATATATGCACTGATGCTTGTATTCCTTGTGTGCTTCTTAAAGTCCCTGAG 1621
Qy 777 ThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796
Db 1622 ACAAAGGGGATGCCCTTGAGGTATTACCGAATTTCTTTCAGTGTGGTGGAAAGACG 1681
Qy 797 AlaAlaLysAla 800
Db 1682 GCTGCAAAAGCC 1693

RESULT 6
US-10-437-963-93472/c
; Sequence 93472, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Roca, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 93472
; LENGTH: 2769
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91853C.1
US-10-437-963-93472

Alignment Scores:
Pred. No.: 1,296-270 Length: 2769
Score: 2835.50 Matches: 562
Percent Similarity: 82.45% Conservative: 77
Best Local Similarity: 72.52% Mismatches: 120
Query Match: 69.58% Indels: 16
DB: 18 Gaps: 7

US-10-051-909-32 (1-800) x US-10-437-963-93472 (1-2769)
Qy 30 AlaLeuProGlyProLeuProProAlaSerCys-----SerSerGl 43
Db 2684 GCCCTTCCAGATTCACAGCGCGCTCTCTCTTTGTTAGGGGATCCGAAATCTCGTGGAGC 2625
Qy 43 nGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuVal 63

```

Db 2624 AGAGACTTGGTGAAGAT-----TCCCGCGGCATGGCGGCGCGTGTGTCGCG 2574  
Qy 63 aileValAlaSerIleGlyAenLeuLeuGlnGlyTrpAspAenAlaThrIleAlaAla 83  
Db 2573 CATCGCGCCTCCATCGCAACTTGTGTCAGGGCTGGGATAATGCAACCACTTCAGGTGC 2514  
Qy 83 aValLeuTyrlleIleIysGlyPheGlnLeuGlnAenGluProThrValGluGlyLeuI 103  
Db 2513 GGTACTGTACATCAAGAAGGAATTCAACTTCGACAGCGAGCCCTTATCGAAGCCCTGAT 2454  
Qy 103 eValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAs 123  
Db 2453 CTTGCCCATGTCGCTCATTTGGGGCGACGATCATCACAGCTTCTCTGACAGATGGCTGA 2394  
Qy 123 pSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyrlPhePheSerGlyLe 143  
Db 2393 TTTCTTTTGGTAGGCGGCCCATGCTGATCGCGTGGCTGTCTCTACTTTGTAGTGGCT 2334  
Qy 143 uIleMetLeuTrpSerProAenValTyrlValLeuLeuLeuAlaAenGlyPheValAspGlyPh 163  
Db 2333 AGTGATGCTTTGGCGCGCAATGTGTATGTGTGCTCTTGGCGAGGCTCATTTGACGGGTT 2274  
Qy 163 eGlyIleGlyLeuAlaValThrLeuValProLeuTyrlIleSerGluIleAlaProSerGl 183  
Db 2273 CGGGATCGGTTTGGCTGTACGCTTGTACCAATGTGTACATCTCTGAGACTGCCCGACGGA 2214  
Qy 183 uIleArgGlyLeuLeuAenThrLeuProGlnPheSerGlySerGlyMetPheLeuSe 203  
Db 2213 CATCAGAGGACTGCTAAACACGCTGCCGCGATTCAGTGGGTCTGAGGGATGTTTCCITTC 2154  
Qy 203 rTyrcyMetValPheGlyMetSerLeuSerProSerProAenTrpArgIleMetLeuGl 223  
Db 2153 ATACTGTAGTGATTTTGGCATGTCCCTCATGCCACAGCCAGATTGGAGGATCATGCTTG 2094  
Qy 223 yValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrlLeuProGluSe 243  
Db 2093 CGTTCTATCAATCACCATCTATATCTTGTGCTTGTGATGACCATCTTTTACTTACCTGAATC 2034  
Qy 243 rProArgTrpLeuValSerIysGlyArgMetAlaGluAlaIysIysValLeuGlnIysLe 263  
Db 2033 GCCGAGGTGGCTCGTGAGCAAGGAAGAAATGGCTGAGGCCAAGCGTGTGTGCAAGCCT 1974  
Qy 263 uArgGlyIysAspValSerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGl 283  
Db 1973 GCGTGGGAAGAGAGATGTTTCAGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGGTGG 1914  
Qy 283 yGlyAspThrSerIleGluGluTyrlleIleGlyProAlaThrGluAlaAlaAspLe 303  
Db 1913 GAAGAACAACAAATTGAGGAATACATAATTGGACCTGATGATGCTGTGATGAAG 1854  
Qy 303 uValThrAspGlyAspIysGluGlnIleThrLeuTyrlGlyProGluGluGlyGlnSerTr 323  
Db 1853 GCTGGCTCCA--GATCCAGAGAAGATCAAACTGTATGGTCTCTGAAGAAGCTTATCGTG 1797  
Qy 323 pIleAlaArgProSerIysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerAr 343  
Db 1796 GGTTCGCCGCTCTTTTACCGGGCAAGTGCACCTTGGAAAGTGCATTAGGTCTCATCTCG 1737  
Qy 343 qHISGlySerMetValAenGlnSerValProLeuMetAspProIleValThrLeuPheGl 363  
Db 1736 TCATGGTAGTATGCTAGTCAGGGTAAGCCCTTGTGGATCCTGTTGTACCCCTTTTGG 1677  
Qy 363 ySerValHISGluAenMetProGlnAlaGlyGlyIysMetArgSerThrLeuPheProAs 383  
Db 1676 AAGTGTCCATGAGAAGATGCTTGAGATAATGGGAAGCATCGGAGACATTTGTTCTCTAA 1617  
Qy 383 nPheGlySerMetPheSerValThrAspGlnHISAlaIysAenGluGlnTrpAspGluGl 403  
Db 1616 CTTTGGCAGCATGTTTGTGTGGCGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGA 1557  
Qy 403 uAenLeuHISArgAspAspGluGluTyrlAlaSerAspGlyAlaGlyIysTyrlGluAs 423  
Db 1556 GAGT---CAACGGGAGGTTGAAGATTATGGATCAGACCATGGTGGGATGACATTGAAGA 1500

Qy 423 pAenLeuHISerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyIysAspIle-- 442  
Db 1499 TAGCCTCCAAAGCCCACTTATTTCTGTCAGGCGACAAGCGTGGAGAAAGGAGATCGC 1440  
Qy 443 -ValHISHisGlyHISArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGl 462  
Db 1439 TGCACCTCTGGCAGTATAATGGTGTGTGGGA-----AGAAGTAGTAGTCTCATGCA 1386  
Qy 462 uGlyIysAspGlyValSerSerThrAspIleGlyGlyTyrlPnGlnLeuAlaTrpIysTr 482  
Db 1385 GGGCGGGGAGCAGTAAGCAGCATGGCATTGTGTGGGGATGGCAGTTGGCTTGGAAATG 1326  
Qy 482 pSerGluIysGluGlyAenGlyArgIysGlyGlyGlyPheIysArgValTyrlLeuH 502  
Db 1325 GACTGAGAGAGAAGTGCAGATGGCGAAAAAAGAGGTGGCTTCCAACGTATCTACTTGCA 1266  
Qy 502 sGlnGluGlyValProGlySerArgGlySerIleValSerLeuProGlyGlyGlyA 522  
Db 1265 TGAAGAGGGTGTGACAGGTGATCGCAGGGGCTCTATCTGTCATTTGCT---GGAGTGA 1209  
Qy 522 pValPheGluGlySerGluPheValHISAlaAlaAlaLeuValSerGlnSerAlaLeuPh 542  
Db 1208 TGTTCCTCTGGTGGTGGTTCGTCAGGAGCTGCTCTGTGTCAGCAACCTGCTCTTTA 1149  
Qy 542 eSerIysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHISProSerGluVa 562  
Db 1148 CTCTAAGGAATTGATGAGCAAGCCTTGTGGCCCTGCTATGCTGTCATCCATCTCAGGC 1089  
Qy 562 lAlaAlaIysGlySerArgTrpIysAspLeuPheGluProGlyValArgArgAlaLeuLo 582  
Db 1088 AGTTGCTTAAGGTCCTAATGGGCAGACTTATTCGAACCTGGAGTGAAGCATGCTCTGTT 1029  
Qy 582 uValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrlTy 602  
Db 1028 TGTGGCATAGGATACAAATCTCTGCAACAGTTTGTGGCNTTAATGGAGTTCTGTACTA 969  
Qy 602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerIysPheGlyLeuSe 622  
Db 968 CACTCCACAAATTTCTTGAGCAAGCTGGTGTGTGTGTTCTTCTTTCGAACAACTTGAGCTAG 909  
Qy 622 rSerAlaSerAlaSerIleIleuIleSerSerLeuThrThrLeuLeuMetLeuProCyH 642  
Db 908 CTCTCATCTGCATCTATTCTTATAGCGGACTGACAACTTCTGATGCTTCCAGCAT 849  
Qy 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrI 662  
Db 848 TGTATTTGCTATGAGGCTCATGATATGTCGGAAGAGGTTTCTTCTCTTCGCAACAT 789  
Qy 662 eProfileLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682  
Db 788 CCCTATCTGTAGTAGCCTAGCTATCTTGAATTTCTGGTCAATATTCTGGATGGGGAC 729  
Qy 682 rLeuAlaHISAlaLeuLeuSerThrValSerValIleValTyrlPheCyScyPheValMe 702  
Db 728 CATGGTTCATGCTCATCTGTCCACAGTCAGTGTGCATCTACTTCTGCTTCTTTGTCAT 669  
Qy 702 tGlyPheGlyProIleProAsnIleLeuCyAlaGluIlePheProThrArgValArgGl 722  
Db 668 GGGGTTTCGGGCTTATTCGAACATTTCTGTGCGAGAGATTTTCCGACCCCTGCTGGG 609  
Qy 722 yLeuCyIleAlaIleCysAlaPheThrPheTrpIleGlyIysAspIleValThrTyrs 742  
Db 608 CATCTGCATAGCCATCTGTGCCCTAACATTTCTGGATCGGTGATATCATTTGTGACATAC 549  
Qy 742 rLeuProValMetLeuAenAlaIleGlyLeuAlaGlyValPheSerIleTyrlAlaValVa 762  
Db 548 CTTCCCGGTGATGCTCAACGCCATTCGACTCGCTGGAGTGTTTTGGAAATCTTACACAGTGT 489  
Qy 762 lCysLeuIleSerPheValPheValPheLeuIysValProGluThrIysGlyMetProLe 782  
Db 488 CTGCATAGGCTTTCTGTTTGTCTTTCATGAAGGTCGCGAGACAAAGGCAATGCTCTCT 429

QY 782 uGluValIleThrGluPheAlaValGlyAlaLysGlnAla 796  
 Db 428 TGAAGTTCATCAGGAGTCTCTCTGTCGGAGCAAGCAGGCC 386

RESULT 7

US-10-425-115-70742  
 ; Sequence 70742, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION: Thomas J.  
 ; APPLICANT: La Rosa, David K.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 70742  
 ; LENGTH: 4330  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_164513C.1  
 US-10-425-115-70742

Alignment Scores:  
 Pred. No.: 1,2e-265 Length: 4330  
 Score: 2788.50 Matches: 538  
 Percent Similarity: 84.23% Conservative: 87  
 Best Local Similarity: 72.51% Mismatches: 110  
 Query Match: 68.43% Indels: 7  
 DB: 18 Gaps: 6

US-10-051-909-32 (1-800) x US-10-425-115-70742 (1-4330)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75  
 Db 769 ATGGGGGCGCGGTATGTCGATCGCGGCTCTATCGCAACTTGTCTGCGAGGCTGG 828  
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGlyGluPheGlnLeuGlnAsn 95  
 Db 829 GACAAATCGCAATTTGCTGAGCGCTCTGTATACATAAAGAAATTCAACTGCGAGAGC 888  
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
 Db 889 GAGCCTCTGATCGAGGCGCTCATCGTCCCATGCTCTATTTGGGCAACAGTCATCAG 948  
 QY 116 ThrPheSerGlyProLeuSerAspSerIleClyArgArgProMetLeuIleLeuSerSer 135  
 Db 949 ACGTTCTCGGGGCGCGCGGCTGCTGTTGGTAGGAGGCCCATGCTGCTCGCTCGGCT 1008  
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
 Db 1009 GTCTCTACTTCGTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068  
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
 Db 1069 CTCGCAAGCTCATTGAGGCTTCGGTATCGTTTGGGCTCACATTTGCTCTCTAC 1128  
 QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
 Db 1129 ATCTCCGAGACTGCGCCGACAGACATTCGTGGGCTGTGTAAACGTTGCGCGATTCAT 1188  
 QY 196 GlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
 Db 1189 GGGTCAGAGGAGATGTTCTCTCTACTGATGCTGTTGGGATGCTCTATGCTCTATG 1248  
 QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
 Db 1249 CCTGATTGAGGCTCATGTTGGAGTTCGTGTCGATCCCGTCACTTATTATTGAGTGG 1308

QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
 Db 1309 ACTGTTCTTCTACTTCCCTGAATCAACAAAGGTGGCTTGTGACCAAGAGGATGCGGAG 1368  
 QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
 Db 1369 GCGAAGAGAGTGTTCGAAAGGCTCGGGGAAGAAGATGTCTCAGGGAGAGTGGCTCTT 1428  
 QY 276 LeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
 Db 1429 CTAGTTGAAGTTCGCGGTAAAGATACACGATATTGAAGAATACATTAATTTGGTCCC 1488  
 QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
 Db 1489 GATGATGAATTCGCTGATGAAGGCTGCTCCA--GATCCAGAGAGATCAAACTATAT 1545  
 QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
 Db 1546 GGAACCTGAAGAAGGCTATCTTGGGTGCCGACCTGTTCGGGGCAAGAAGTCTCTTGA 1605  
 QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMet---ValAsnGlnSerValProLeu 354  
 Db 1606 AGCGCTTAGTCTCATCTCTCTCATGGAGTATGCGGCTAGTCAGGGTAAGCCCTC 1665  
 QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374  
 Db 1666 GTGGATCTTATGTCCTCTCTTCGAAAGTGTTCATGAAGAAGATGCCTGAGATCATGGG 1725  
 QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394  
 Db 1726 AGCATGAGGAGCACAATGTTTCCCAACTTGGCAGCATGTTTAGTGTTCGCCACCGCAG 1785  
 QY 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414  
 Db 1786 CAGGTCAAAAGCTGACTGGGACCGCAGAGT---CAAAGGAAAGGTGAAGATTATCTTCG 1842  
 QY 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434  
 Db 1843 GATCATGGTGGCGATGACATTCAGGATTAACCTCCAAAGGCCACTTATTCTCGTCAGGCA 1902  
 QY 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454  
 Db 1903 ACAAGTGTGAAGAAAGAGAGATCGCT-----GCACCTCATGCTAGCATATTGGGTGCT 1956  
 QY 455 ArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGlyGly 474  
 Db 1957 GTGGGAAGGAGCAGTAGCTTCGAGGAGGGGAGGAGGAGTAAAGCAGCATGGCCATTGGCGGA 2016  
 QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlyGluAsnGlyArgLysGlyGly 494  
 Db 2017 GGATGGCAGTTGGCGTGAATGACCCGAGAGAGGCGAGAGTGGGCAAGAAGAGT 2076  
 QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514  
 Db 2077 GGCTTCCAGCGTATTACTTTGATGAGGAGGGGTACAAGGCCAAC---AGGGTCTATA 2133  
 QY 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 534  
 Db 2134 TTGTTCATTACCA---GGCGGGGATGTTCTCTCTGTTGGTGGTTCATCCAGGCTCAGCT 2190  
 QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554  
 Db 2191 CTTGTGAGCCAAACCCAGCTCTTTACTCTTAAGGAACCTCTCGAGCAACGCTGCTGCTCT 2250  
 QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574  
 Db 2251 GCGATGATGATCCATCTGAAGCAGTTCCTAAAGGTCCAAGATGGGCGCCGCTATTGAG 2310  
 QY 575 ProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594  
 Db 2311 CTTGGGTGAAGCATGCTGCTGTTTGGCATAGGAATACAGATCTCTGCAACAGTTTGGT 2370  
 QY 595 GlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614

Db	2371	GGCATCAACGGCGTTCCTACTACATCCCTCAAAATCTTGAGCAAGCAGCGTCGGGTGT	2430
Qy	615	IleLeuSerIyPbPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr	634
Db	2431	CTTCTGTGCAACCTCGGCGCTTAAACGCTTCTTCGCGCATCAATCCTCATTAGCGCCCTCAGC	2490
Qy	635	ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg	654
Db	2491	ACCTTACTGATGCTCCCAAGCATCGGCATTCGCATGAGGCTCATGGATATGTTCGGGAAGG	2550
Qy	655	ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal	674
Db	2551	AGGTTTCTCTCTCGCGACGATCCAGTCCTAATAGTCGCGCTACTCGTCTCGTGTGTG	2610
Qy	675	SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle	694
Db	2611	TCCAACATCGTGGACGTGGGGACGTGGCGACCGCGCGTCTCTCCACGGCAGCGTCATA	2670
Qy	695	ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu	714
Db	2671	GTCTACTTCTGCTTCTTCGTCAUGGGGTTCGGGCCGCTCCCAACATCCTCTCGCGCAGAG	2730
Qy	715	IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle	734
Db	2731	ATCTTCCCCACACGGTCCGCGGTGTGTGTCATCGCCATCTCGCGCCTTCGTGGGCT	2790
Qy	735	GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly	754
Db	2791	GGTGACATCATCGTAGCATCCTCTCCCGTGATGCTGAACTGCTCGGGCTCGCCGCC	2850
Qy	755	ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuIysVal	774
Db	2851	GTCTTTGGGGTGTACGCGCGTCGTGTGGGTCTCTAGCCCTCGCGTTCGTGTGTTCGTCAAGGTG	2910
Qy	775	ProGluThrIySgIyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaIys	794
Db	2911	CCCGAGCAGAGGCGATCGCTCTCGAGGTATACACGAGTTCCTCTCCGTTGGGGCAAG	2970
Qy	795	GlnAla	796
Db	2971	CAAGCC	2976

RESULT 8  
US-10-051-902-7  
; Sequence 7, Application US/10051902  
; Publication No. US20020178468A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/10/051,902  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97

Alignment Scores:		
Pred. No.:	1,286-254	2601
Score:	2674.00	522
Percent Similarity:	82.20%	Conservative: 115
Best Local Similarity:	69.88%	Mismatches: 92
Query Match:	65.62%	Indels: 18
DB:	13	Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-902-7 (1-2601)

Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAenLeuLeuGlnGlyTrp	75
Db	175	ATGAAGGTGCGCTCTTGTTGCTATTGCGCGCTCCATTGGTAATTTCCTCCAAGGATGG	234
Qy	76	AspAsnAlaThrIleAlaAlaValLeuTyrlleLysLysGluPheGlnLeuGlnAsn	95
Db	235	GATAATGCTACCATCGCGGGCTTAATGGTTACATTAAAGAAGACCTTGCTTTGGGAACA	294
Qy	96	GluProThrValGluGlyIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	295	-----ACTATGAAAGGCTTGTTGGTGGGCATGTCCTGATTGGAGCAACGGTAATCACC	348
Qy	116	ThrPheSerGlyProLeuSerAspSerIleGlyValArgProMetLeuIleLeuSerSer	135
Db	349	ACATGCTCTGGTCTATAGCGGATTGGCTCGGTGCGGACCCATGATGATTAATCTCATCT	408
Qy	136	IleLeuTyrlPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrlValLeuLeu	155
Db	409	GTGCTCTATTTCTTGGTGGTTTGGTGATGCTGTGGTCCCAATGTGTATGTTGTGTGC	468
Qy	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrl	175
Db	469	TTGCGGAGGCTACTTGATGGAATTTGGGATTTGGCTCTGCTGATCTCTTGTCGCGTCTAT	528
Qy	176	IleSerGluIleAlaProSerGluIleArgGlyLeuAenThrLeuProGlnPheSer	195
Db	529	ATATCTGAACGGCGCGCTCTGAATAAGGGGTGCTGTAATACGCTTCCTCAGTTCCAGT	588
Qy	196	GlySerGlyGlyMetPheLeuSerTyrlCysMetValPheGlyMetSerLeuSerProSer	215
Db	589	GGCTCTGGAGGAATGTTTGTGCTACTGTATGGTTTGGCATGTCATTGAGTCCCGCG	648
Qy	216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerIleuPhePheGlyLeu	235
Db	649	CCTAGCTGGAGGCTCATGCTTGGGGTCTGCTATTCTCTCTCTCTGTAATTTGCATTG	708
Qy	236	ThrIlePheTyrlLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255
Db	709	ACCATTTTTTCTTGCCCGAGTCTCTCGGTGGCTGCTGACGAAGAAGATGCTCGAG	768
Qy	256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu	275
Db	769	GCTAAGAGGTGCTCCAAAGATTGCGCGAAGGAGGATGCTGTGAGGCGAGATGGCATTG	828
Qy	276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrlleIleGlyPro	295
Db	829	CTGTTGMAAGTCTCGGATTTGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT	888
Qy	296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrl	315
Db	889	GCTGACGATGTGCTGATGTCATGAACATGCACACAGAGAAGATAAAATTCGATTATAT	948
Qy	316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335
Db	949	GGATCCCAAGCAGCCCTTCTTGTTATCAAAACCTGCTCACTGGACAGAGTTCTATTGGC	1008
Qy	336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Db	1009	-----CTTGCTCACACCATGGAAGCATCATCAACCAAGCATGCGCCCTCATG	1056
Qy	356	AspProIleValThrIleuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
Db	1057	GATCCCTCTGGTGACACTGTTTGGTAGCATTTCAATGAAGAGCTCCCGGACAGACGAAGA	1116
Qy	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1117	GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGGAAGCATGTTTCAGCACTGCTGAGCG	1176
Qy	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnHisArgAspAspGluGluTyrlAla	413

```
Db 1177 CATGCTAAATTTGAACAATGGGATCAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG 1236
QY 414 SerAspGlyAlaGlyValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaVal 791
Db 1237 TCAGATGCAACCGGTGGGACTCCGATGATAATTTGCACAGTCTTTAAATCTCACGCCAA 1296
QY 434 AlaThrGlyAlaGluGlyLeuValIleHisGlyHisArgGlySerAlaLeu--- 452
Db 1297 ACAACAGCCTTGAA---AAAGCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1353
QY 453 SerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 471
Db 1354 AGCATGAGCGGTCAAGTAGTCTCATGCAAGGCTCAGGTGAGCAAGGTGGTGTGTACAGGT 1413
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db 1414 ATTGGTGTGGCTGCACTGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGG 1470
QY 492 LysGluGlyGlyPheLysArgValTrpLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1471 CAACAGGAGGGTTTAAAGAGTATTTATTTACATGAGGAGGAGTCTTCTGCACTCTCGTCGT 1530
QY 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
Db 1531 GGATCCATTGTATTCGATCCCGGTGAAGGC-----GAATTTGTCCAG 1572
QY 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db 1573 GCTGTGCTCTTGGTAAGCCACCCGCTCTTTACTCCAGGAGCTTATTTGATGGACACCCA 1632
QY 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
Db 1633 GTTGGGCTCTGCAATGGTTCCACCATCTGAGACAGCTTCAAAGGGGCCAAGTTGGAAGCT 1692
QY 572 LeuPheGluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTTCTTTGAACAGGGGTAAAGCATGCTGGTGTGGAGTTGGAATACAAATATCTTCAG 1752
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrrTrpProGlnIleLeuGluGlnAlaGly 611
Db 1753 CAGTTTTCAGGGATAAATGGGTTCTATATTACACACCTCAATCTTTGAAGGCCCGGT 1812
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTTCTCTTTTCAGATATAGGCATTTGGCTCAGAGTCGGCATCATCTCTTATCAGT 1872
QY 632 SerLeuThrThrLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTCAGAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCATGAAGCTCATGATGTT 1932
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAGAGGCGAGTGTCTACTTACTACATCCCGGTGCTGATTTGCTCACTCATATT 1992
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGGTCATTTGGAAGCCTGTGTAATTTGGCAATGTGCGCCATGCGCAATCTCAACAGTA 2052
QY 692 SerValIleValTrpPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TGGTGTGTGTTATTCTGTGCTGTGTGTGATGGGTATTTAGGACCAATTTCCAAACATCCT 2112
QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TGCTCAGAGATTTTCCCCACTAGGTGGCTGCTGCTCTGCTATGCTATCTGTGCTATTAGT 2172
QY 732 PheTrpIleGlyAspIleIleValThrTrpSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCTGATTTGGAGACATCATCATCACTACTCGCTGCGTGTGATGCTCGGCTCTTTAGGA 2232
QY 752 LeuAlaGlyValPheSerIleTrpAlaValValCysLeuIleSerPheValPhe 771
Db 2233 CTTGGTGGTGTATTCGCACTTTACGAGTTGTGTGTTCACTCTCGTGATATTGTGTTT 2292
```

```
QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaVal 791
Db 2293 TTGAAGGTTTCAGAAACAAGGCGATGCCCTTGAAGTCATCTCTGAATTTCTTTCTGTT 2352
QY 792 GlyAlaLysGlnAlaAlaAla 798
Db 2353 GGAGCAAGACGAGGTGCTTCT 2373
RESULT 9
US-10-051-909-7
; Sequence 7: Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: HelentJaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-909-7
Alignment Scores:
Pred. No.: 1,288-254 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: 13 Gaps: 8
US-10-051-909-32 (1-800) x US-10-051-909-7 (1-2601)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 175 ATGAAGGTGGCGTCTTGTGTGTTATGCCGCTTCCATTTGGTAATTTCTCTCAAGGATGG 234
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrrIleLysLysGluPheGlnLeuGlnAsn 95
Db 235 GATAATGCTACCATCCCGGGCTAATGGTTACATTAAGAAAGACCTTGTCTTGGGAACA 294
QY 96 GluProThrValGluGlyLeuValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGAAAGCTTGTGTGGGCATGTCCTGATTGGAGCAACGGTAATCAC 348
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 349 ACATGCTCTGGTTCCTATAGCGATTTGGCTCGGTGGCGACCATGATGATAATCTCATCT 408
QY 136 IleLeuTyrrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTrpValLeuLeu 155
Db 409 GTGCTCTATTTCTGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 468
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrr 175
Db 469 TTGGCGAGGCTACTGATGATTTGGGATTTGGGATTTGGGCTTGTGTGACTCTTGTCCGGTCTAT 528
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 529 ATATCTGAAACCGCGCGCTCTGAAATAAGGGGGTGTGGTGTGGTGTGGTGTGGTGTGG 588
QY 196 GlySerGlyGlyMetPheLeuSerTyrrCysMetValPheGlyMetSerLeuSerProSer 215
```



589 GGCTCTGGAGGAATGTTTTTGTCTGTA CTGTATGTTGTTTTTGGCAGTGCATTTAGTGTCCGCG 648

216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235

649 CCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCTCTCTCTGTATTTCATTC 708

236 ThrIlePheTyrlLeuProGluSerProArgTrpLeuValSerIysGlyArgMetAlaGlu 255

709 ACCATTTTTTCTTGCCCGAGTCTCTTCGGTGGCTGTACGAAAGAGGATGCTCGAG 768

256 AlaIysIysValLeuGlnIysLeuAqtGlyAspAspValSerGlyGluLeuSerLeu 275

769 GCTAAGAAGGTGCTCCAAGATTCGCGGAAGGAGGATGTGTACGCGAGATGGCATTC 828

276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluIuTyrlIleIleGlyPro 295

829 CTGTTGAAGTCTCGGGAATGGGGGTGATACATCTATCCGAAGTACATAAATGGCCCT 888

296 AlaThrGluAlaIleAspLeuValThrAspGlyAspIysGluGlnIleThrLeuTyrl 315

889 GCTGACGATGTGGTGTATGTCATGAACATGCACAGAGAAGATAAAATTCGATTATAT 948

316 GlyProGluGluGlnSerTrpIleAlaArgProSerIysGlyProIleMetLeuGly 335

949 GGATCCCAAGAGCGCTTTCTTGTTTATCAAAACCTGTCTACGTGACAGATTTCTATTGGC 1008

336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355

1009 -----CTTGGCTCACACCTGGAAAGCATCATCAACCAAGCATGCCCTCATG 1056

356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373

1057 GATCCTCTGCTGACACTGTTTGGTAGCATTCATGAGAAGCTCCCGAGACAGAGCAAGA 1116

374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393

1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGGAAAGCATGTTCAGCACTGCTGAGCGC 1176

394 HisAlaIysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluIuTyrlAla 413

1177 CATGCTAAATTTGNAACAAATGGGATGAAGAAGCTTACAAGGGNACGTGAGGACTACATG 1236

414 SerAspGlyAlaGlyGlyAspTyrlGluAspAsnLeuHisSerProLeuLeuSerArgGln 433

1237 TCAGATCAACCCGTGGGGACTCCGATGATAATTTTGCACAGTCTCTTAAATCTCACGCCAA 1296

434 AlaThrGlyAlaGluGlyIysAspIleValHisGlyHisArgGlySerAlaLeu--- 452

1297 ACAACAGCCCTTGAA---AAAGACTTACCTCTCTCTCTCCCATGGCAATATCTCTTGGC 1353

453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471

1354 AGCATGAGGGCTACAGTAGTCTCATCAAGGTCAGGTGAGCAGAGGTGTGTAGTACAGGT 1413

472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluIysGluGlyGluAsnGlyArg 491

1414 ATTGGTGGTGGCTGGCAACTGGCATGGAATGACATGATAAA---GGTAGGATGGAAA 1470

492 LysGluGlyGlyPheIysArgValTyrlLeuHisGlnGluGlyValProGlySerArgArg 511

1471 CAACAGAGGGGTTTAAAGGATTTATTTATCATGAGGAGGGAGTTTCTGCATCTCTCGT 1530

512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531

1531 GGATCCATTTGTATTCGATTCCCGTGAAGGC-----GAATTTGTCAG 1572

532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerIysGlyLeuAlaGluProArgMet 551

1573 GCTGCTGCTTGGTAAGCAACCGCTCTTTACTTCCAAGGAGCTTATTGTATGGACACCCA 1632

552 SerAspAlaAlaMetValHisProSerGluValAlaAlaIysGlySerArgTrpLysAsp 571

1633 GTTGGGCGCTGCAATGGTTCCCATCTGAGACAGCTTCAAGGGGGCCAAAGTGTGGAAGCT 1692

```

1  RESULT 10
2  US-09-938-842A-1315
3  ; Sequence 1315, Application US/09938842A
4  ; Patent No. US20020160378A1
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Harper, Jeff
7  ; APPLICANT: Kreps, Joel
8  ; APPLICANT: Wang, Xun
9  ; APPLICANT: Zhu, Tong
10 ; TITLE OF INVENTION: STRESS-REGULATED G
11 ; TITLE OF INVENTION: SAME, AND METHODS
12 ; FILE REFERENCE: SCIP1300-3
13 ; CURRENT APPLICATION NUMBER: US/09/938,
14 ; CURRENT FILING DATE: 2001-08-24
15 ; PRIOR APPLICATION NUMBER: US 60/227,86
16 ; PRIOR FILING DATE: 2000-08-24
17 ; PRIOR APPLICATION NUMBER: US 60/264,64
18 ; PRIOR FILING DATE: 2001-01-16
19 ; PRIOR APPLICATION NUMBER: US 60/300,11
20 ; PRIOR FILING DATE: 2001-06-22
21 ; NUMBER OF SEQ ID NOS: 5379
22 ; SEQ ID NO 1315
23 ; LENGTH: 2190
24 ; TYPE: DNA

```

RESULT 10  
 US-09-938-842A-1315  
 ; Sequence 1315, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE  
 ; FILE REFERENCE: SCRIPT100-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 1315  
 ; LENGTH: 2190  
 ; TYPE: DNA

; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1315

Alignment Scores:	
Pred. No.:	8,488-246
Score:	2584.00
Percent Similarity:	60.78%
Best Local Similarity:	68.92%
Query Match:	63.41%
DB:	9
Length:	2190
Matches:	512
Conservative:	89
Mismatches:	123
Indels:	20
Gaps:	8

US-10-051-909-32 (1-800) x US-09-938-842A-1315 (1-2190)

Qy	56	MetSerGlyAlaValLeuValAlaValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	1	ATGAGTGGAGCTGTGCTGTGTCTATGCTGCTGCTGTGTGGCACTGTTCACGAGATGG	60
Qy	76	AspAsnAlaThrIleAlaAlaValLeuTyrlleIlylsyGluPheGlnLeuGlnAen	95
Db	61	GATAACGCACATATTTCGAGGAGCTGTGTGTACATAAAAAAGAGATTAAATTTGGAGAGT	120
Qy	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	121	AAATCCATCAGTCGGAAGGCTTAATTTGGCGGATGTCACTTATTTGGTGCTACTCTGATTAACA	180
Qy	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Db	181	ACATGCTCTGGAGGGGTAGCTGATTGGCTGCCCGTCCCATGTGTAATATTGTCCTCA	240
Qy	136	IleLeuTyrlPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrlValLeuLeu	155
Db	241	ATTCCTACTTTGTTGGTCTCTAGTAATGCTATGGTCTCCGAATGTTATGTGTGCTC	300
Qy	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrl	175
Db	301	TTAGGAAGTTGTTAGATGGATTGGGGTTGGTCTTGTGGTCACACTTGTCTCTATTAT	360
Qy	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Db	361	ATATCTGAGACTGCACCACTGAGATATAGGGGACTGTGAATAGCTACGCGAGTTCAC	420
Qy	196	GlySerGlyGlyMetPheLeuSerTyrlCysMetValPheGlyMetSerLeuSerProSer	215
Db	421	GGCTCTGGAGGAGATGTCCTATCTTACTGTATGTGTTTCGGAATGCTGTGATGCCATCA	480
Qy	216	ProAspTrpArgIleMetHeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Db	481	CCTAGCTGGAGATGTATGCTGCTGCTCTTTTCATCTCCCTTGCTTTTCTTCCTC	540
Qy	236	ThrIlePheTyrlLeuProGluSerProArgTrpLeuValSerIysGlyArgMetAlaGlu	255
Db	541	ACGGTCTCTCTTGTCCCGAGTCCCAAGGTGGCTGCTGAGCAAAAGTCGAATGCTTGAA	600
Qy	256	AlaIysAluValLeuGlnIlyLeuArgGlyIysAspValSerGlyGluLeuSerLeu	275
Db	601	GCAAAGCGGGTCTTCAGAGACTGCGTGGTCGCAAGATGTGTGTGTGAGATGGCTTGT	660
Qy	276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrlleIleGlyPro	295
Db	661	TTGTTTGGAGGCTTCGAAATTTGGATGGAGTGAACACCATAGNGGNATATATAATTTG	720
Qy	296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIysGluGlnIleThrLeuTyrl	315
Db	721	GCGGATGAAGTTACTGATGATCATGATATAGCTGTGGATATAGGATCAATTAAGTTAT	780
Qy	316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerIysGlyProIleMetLeuGly	335
Db	781	GGTGCAAGAAAGGGCTGAGTTGGGTTGCTAGCGCAGTCAAAAGA-----GGA	828
Qy	336	SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu	354
Db	829	AGCACTATGAGTTGTGTTGCTCGCAATGGGAAGTACAAATGACAGGAGGAGGACTCATTTG	888

355	Met	App	Pro	Ile	Val	Thr	Leu	Phe	Gly	Ser	Val	His	Glu	Asn	Met	Pro	Gln	Ala	Gly	374	
889	AT	TGA	CT	CT	GT	GT	CA	CA	CT	GT	TT	GG	AG	CG	GT	TT	CA	GAG	AG	AT	945
375	Ser	Met	Arg	Ser	Thr	Leu	Phe	Pro	Asn	Phe	Gly	Ser	Met	Phe	Ser	Val	Thr	Asp	Gln	His	394
946	AG	CA	TGA	GAG	AG	TG	CT	TT	CC	CA	CA	TT	TT	GG	AG	TAT	GT	TC	AG	GA	1005
395	Ala	Lys	Asn	Glu	Gln	Trp	Asp	Glu	Leu	Asn	Leu	His	Arg	Asp	Asp	Glu	Glu	Tyr	Ala	Ser	414
1006	CA	GA	CA	TGA	CA	GA	AG	AT	TC	GG	AT	GA	AG	AT	CT	TT	GG	AG	AG	GT	1065
415	Asp	Gly	Ala	Gly	Gly	Asp	Tyr	Glu	Asp	Asn	Leu	His	Ser	Pro	Leu	Leu	Ser	Arg	Gln	Ala	434
1066	GAC	---	CAT	GAG	AT	GAT	CT	GA	AG	AT	GAT	CT	CA	TT	CT	CG	GT	AT	CT	CA	1122
435	Thr	Gly	Ala	Glu	Gly	Lys	Asp	Ile	Val	His	His	Gly	His	Arg	Gly	Ser	Ala	Leu	Ser	Met	454
1123	ACA	AG	CA	TGG	AG	---	AAA	GA	CA	TGC	CT	CA	CT	GC	TT	CA	TGG	AA	CT	TT	1179
455	Arg	Arg	Gln	Thr	Leu	Leu	Gly	Glu	Gly	Gly	Arg	Gly	Val	Ser	Ser	Thr	Asp	Ile	Gly	Gly	474
1180	CAT	GAA	AG	TCA	AG	TGC	AG	CG	CT	CA	AG	GG	CA	AG	CG	GG	TAT	GAT	GG	AT	1239
475	Gly	Trp	Gln	Leu	Ala	Tp	Lys	Trp	Ser	Glu	Lys	Glu	Gly	Val	Asn	Gly	Arg	Lys	Glu	Gly	494
1240	CGA	TGC	CA	AG	TGG	CA	TG	GA	AT	CG	CA	AA	GA	GA	AT	GA	TG	CG	GA	CA	1296
495	Gly	Phe	Lys	Arg	Val	Tyr	Leu	His	Gln	Glu	Gly	Val	Pro	Gly	Ser	Arg	Arg	Gly	Ser	Ile	514
1297	-----	-----	-----	-----	-----	-----	GA	AG	TT	TT	CC	CA	GG	AT	CT	CG	AG	CG	TT	CA	1329
515	Val	Ser	Leu	Pro	Gly	Gly	Gly	Asp	Val	Phe	Glu	Gly	Ser	Glu	Phe	Val	His	Ala	Ala	Ala	534
1330	GT	TT	CA	TG	CT	CT	GT	GT	GT	AT	GA	CG	GT	AG	---	GC	AG	AT	TT	GT	1386
535	Leu	Val	Ser	Gln	Ser	Ala	Leu	Phe	Ser	Lys	Gly	Leu	Ala	Glu	Pro	Arg	Met	Ser	Asp	Ala	554
1387	TTG	GT	TAT	GC	CA	AC	CAG	CT	CT	TAT	CA	AA	GA	CT	CT	CA	AA	GA	CA	TCA	1446
555	Ala	Met	Val	His	Pro	Ser	Glu	Val	Ala	Ala	Lys	Gly	Ser	Arg	Trp	Lys	Asp	Leu	Phe	Glu	574
1447	GC	TAT	GT	GT	TAC	AT	CA	CT	CG	AA	---	ACA	ACT	TAA	AGG	GT	CA	AT	TT	GG	1503
575	Pro	Gly	Val	Arg	Arg	Ala	Leu	Leu	Val	Gly	Val	Gly	Ile	Gln	Ile	Leu	Gln	Ala	Phe	Ala	594
1504	CCT	GAG	TCA	AG	CG	TG	CA	TT	AG	TG	CT	AG	AG	TG	TT	GG	CA	TAT	CT	CA	1563
595	Gly	Ile	Asn	Gly	Val	Leu	Tyr	Thr	Pro	Gln	Ile	Leu	Glu	Gln	Ala	Gly	Val	Ala	Val	614	
1564	GG	CA	TCA	CG	CG	AG	TT	CT	TT	ACT	CA	CA	CG	CA	AA	TCT	TT	GAG	CG	CG	1623
615	Ile	Leu	Ser	Lys	Phe	Gly	Leu	Ser	Ser	Ala	Ser	Ile	Ile	Leu	Ile	Ser	Ser	Leu	Thr	634	
1624	CT	AT	CT	CG	AA	CT	GG	AT	TAG	TT	CT	CT	CT	CA	GC	AT	CT	CT	TAT	TAA	1683
635	Thr	Leu	Met	Leu																	

```

Db      1924  ATTTTCCAACTCGAGTCCCGGAATCTGCATCGCCATCTGCCACTCACCTTCTGGATC 1983
Qy      735  GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
Db      1984  TGTGACATAATCGTCACTTACAGCTCCCGCTGCTGCTCAATCCATGGACTAGCTGGT 2043
Qy      755  ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774
Db      2044  GTGTTTGGAAATGACCAATCGATGTGCAATTTCAATGGGTCTTGTGTTCATTAAGTC 2103
Qy      775  ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
Db      2104  CCGAAACTTAAGGCATGCCACTTGAAGTCATCAGAGTTCCTTCTGTGTGGAGCTAGA 2163
Qy      795  GlnAlaAlaAla 798
Db      2164  CAAGCTGAAGCT 2175

RESULT 11
US-09-938-842A-1315
; Sequence 1315, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1315
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1315

Alignment Scores:
Pred. No.:      8,48e-246      Length:      2190
Score:          2584.00      Matches:      512
Percent Similarity: 80.78%      Conservative: 89
Best Local Similarity: 68.82%      Mismatches: 123
Query Match:     63.41%      Indels:      20
DB:              11          Gaps:         8

US-10-051-909-32 (1-800) x US-09-938-842A-1315 (1-2190)
Qy      56  MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db      1  ATGAGTGGAGCTGCTGTTGTTGCTATTGCTGCTGCTGTTGGCAACTTGTACAGGATGG 60
Qy      76  AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGlyGluPheGlnLeuGlnAsn 95
Db      61  GATAACGCCAACTATTTCAGGAGCTGTGTGTGTACATAAAAAGGAGTTTAATTTGGAGAGT 120
Qy      96  GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db      121  AATCCATCAGTGAAGGCTCAATTGTGGCGATGCTCACTTATGCTACTCTGTATTACA 180
Qy      116  ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db      181  ACATGCTCTGGAGGGTAGCTGATTGGCTGGTGGCTGCCGTCCTCATGCTATATTGCTCTCA 240
Qy      136  IleLeuTyrPhePheSerGlyLeuIleMetLeuLeuTrpSerProAsnValTyrValLeuLeu 155

```

```

Db      241  ATTCTCTACTTGTGTGGTCTCTAGTAATAGCTATAGTCTCCGATGTTTATGTGTGCTC 300
Qy      156  LeuAlaAraPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db      301  TTAGGAAGGTTGTAGATGGATTGGGTGGTGGTCTTGTGGTCACACTTGTTCCTATTAT 360
Qy      176  IleSerGluIleAlaProSerGluIleAraGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db      361  ATATCTGAGACTGCACACCTGAGATTAGGGGACTCTTGAATACGTCACGTCAGTTCAC 420
Qy      196  GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db      421  GGCTCTGGAGGATGTTCTTATCTTACTGTATGGTGTTCGGAATGCTGTTGATGCCATCA 480
Qy      216  ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerSerLeuPhePheGlyLeu 235
Db      481  CCTAGCTGGAGATTGATGCTGTGTGCTTTTCATCCCTCCCTCTCTTCTTCTTCTCTC 540
Qy      236  ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db      541  ACGGCTTCTTCTTGTCCCGAGTCCCAAGGTGGCTCGTGAGCAAAAGGTGGAATGCTTGA 600
Qy      256  AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275
Db      601  GCAAAGCGGGTCTTCAGAGACTGCGTGGTGGCAAGATGTGTCTGTGTAGATGGCTTGG 660
Qy      276  LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleLeuGlyPro 295
Db      661  TTGGTTGAGGGTCTTGGAAATGGAGGTGAACAACCATAGAGGAATATATAATTTGTTCC 720
Qy      296  AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db      721  GCGGATGAAGTTACTGATCATGATATAGCTGTGGATAAGGATCAAAATTAAGTTATAT 780
Qy      316  GlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db      781  GGTGCAGAAAGAGGCTGAGTTGGTGTCTAGCCAGTCAAAAGG-----GGA 828
Qy      336  SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354
Db      829  AGCACTATGAGTGTGTTTGTCTGCCATGGAAGTACAATGAGCAGGAGGCAAGGCTCAT 888
Qy      355  MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
Db      889  ATTGATCCTCTGTGCACACTGTTGGAGCGTTTACGAGAGATGTCGCGACACT---GGA 945
Qy      375  SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
Db      946  AGCATGAGGAGTGCCCTTGTTCACATTTTGGAGTATGTTCAAGTTTGGAGGGAATCAA 1005
Qy      395  AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGluTyrAlaSer 414
Db      1006  CCAAGACATGAAGATTGGGATGAAGAGATCTTGTGGAGAGGTTGAGGATTATCCATCC 1065
Qy      415  AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db      1066  GAC---CATGGAGATGATCTGAAGATGATCTTCAATCTCCGTTGATCTCACGTCAAACG 1122
Qy      435  ThrGlyAlaGluGlyLysPheIleValHisGlyHisArgGlySerAlaLeuSerMet 454
Db      1123  ACAAGCATGGAG---AAAGACATGCTCCTCACCTGCTCATGGAACTCTTCTACCTTCAGA 1179
Qy      455  ArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
Db      1180  CATGGAAGTCAAGTCAGGAGGCTCAAGGGGAAGAGCGGATAGTATGGGATTTGAGGT 1239
Qy      475  GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494
Db      1240  CGATGCAAGTGGCATGGAAATGACCGAAAGAGAGATGAATCGCGACAGAAAGAA--- 1296
Qy      495  GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIle 514

```

```

Db 1297 -----GAAGTTTCCAGGATCTCGAGCTGGCTCAATT 1329
QY 515 ValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAala 534
Db 1330 GTTTCATTGCTGGTGGTGAAGACCGGTAG--GCAGATTGTACAGGCTGCT 1386
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
Db 1387 TTGGTTAGCAACACGAGCTCTTATTCCAAAGACCTTCTCAAAGAACATACATAATGGTCCCT 1446
QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgfipLysAspLeuPheGlu 574
Db 1447 GCTATGGTACATCCATCCGAA---ACAACATAAGGGTCAATTTGCGCATGATCTTCATGAT 1503
QY 575 ProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594
Db 1504 CTTGGAGTCAACGGTGCATTAGTCTAGGAGTTGGACTTCAATACATTCAGCAGTCTCA 1563
QY 595 GlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614
Db 1564 GGCATTCACGGAGTCTTTACTACACCGCAATCTTTGAGCGGGGTGTCGGGATC 1623
QY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleLeuSerSerLeuThr 634
Db 1624 CTACTATCGAATGGGAGTAGTCTTCTTCAGCATCTTACTTATAAGTGATGATCA 1683
QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
Db 1684 ACCTTTGTGATGTTACCTGCAATAGCTGTGCAATGAGGCTCATGATCTTTCTGGTCA 1743
QY 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
Db 1744 AGGACCTTGCTTCTCACCACGATACCAATCTCTGATAGCATCTTATTGGTTTATGTAATC 1803
QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
Db 1804 TCAATCTTGTTCACATGAACAGCATGTGTCACCGGGCTTTATCAACCGTAAAGCGTTGTG 1863
QY 695 ValTyrPheCysCysPheValMetPheGlyProIleProAsnIleLeuCysAlaGlu 714
Db 1864 CTCCTACTCTGCTTCTCGTATGGTTCGGTCTCGTCTGCTCCCAACATCCTCTGTTACAG 1923
QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheThrIle 734
Db 1924 ATTTTTCACATCGAGTCCCGGATCTGCATCGCATCTGGCATCTCACCTTCTGGATC 1983
QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleClyLeuAlaGly 754
Db 1984 TGTGACATAATCGTCACTTACAGTCTCCCGGTGCTGCTCAATTCATTTGGACTAGCTGGT 2043
QY 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774
Db 2044 GTGTTTGGAAATGACGCAATCGTATGTTGCAATTCATGGGTCTTTGTGTTCAATTAAGTC 2103
QY 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
Db 2104 CCGGAACATAAGGCATGCCACTTAAGTCAATCACAGAGTCTCTTTCTGTTGGAGCTAGA 2163
QY 795 GlnAlaAlaAla 798
Db 2164 CAAGCTGAAGCT 2175

```

RESULT 12

```

US-10-424-599-10324
; Sequence 10324, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

```

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 10324
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109333C.1
; US-10-424-599-10324

Alignment Scores:
Pred. No.: 2,266-244 Length: 2671
Score: 2571.00 Matches: 517
Percent Similarity: 80.73% Conservative: 99
Best Local Similarity: 67.76% Mismatches: 107
Query Match: 63.09% Indels: 40
DB: 17 Gaps: 12

US-10-051-909-32 (1-800) x US-10-424-599-10324 (1-2671)
QY 55 LysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGly 74
Db 244 GAAATGAAAGGTACCGTCTCGTGGTATTCGCGTATTCGCGCAATATCTCCAGGGA 303
QY 75 TrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnGln 94
Db 304 TGGGATATGCTTCCATCGCGGGGCCATTGTTTACATTAAAGAAAGACCTTGGCGTTGCAA 363
QY 95 AnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 364 ACA-----ACTATGGAAGGCTGTGTGGCCATGCTCCCTGATTTGGACCAACGGTAATC 417
QY 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuLeuSer 134
Db 418 ACCACATGCTCTGGTCTTATAGCGGATTTGGCTCGGCGACCATGATGATATCTCA 477
QY 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154
Db 478 TCTGTGCTCTATTTCTTGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 537
QY 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 538 TGTCTGGCGAGGCTACTTGTGATGATTTGGGATTTGGGCTTGTGTGACTCTTTGTCCCGGTC 597
QY 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 TATATATCTGAAACGCGCGCTCTCAAAATAAGGGGGTCTGTTGAATACGCTTCTCCAGGTC 657
QY 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 658 AGTGGCTCTGGAGGAATGTTTTCGTACTTATGTTTGGTATGTTTGGCATGCTCATTGAGTCCC 717
QY 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
Db 718 GCGCCTAGCTGGAGGCTCATGCTTGGGGTCTGTCTATTCCTTCCCTTCTGTTATTTTGA 777
QY 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet--- 253
Db 778 TTGACCAATTTTCTTCTGCGCGAGTCTTCC-----TCTTCAGGCTCGGTGAGC 825
QY 254 -----AlaGluAlaLysIstysValLeu-GlnLysLeuArgGlyLysAs 267
Db 826 AAAATGAAATTTGATGCTCGGACGCGCAACAAAGTGTCTCCAAAGATTTGCGCGGACGGA 885
QY 267 pAspValSerGlyGluLeuSerLeuLeuGluGly-LeuGlu-ValGlyGlyAspThr 286
Db 886 GGATGTGCTGGCGAGATGACCTTGTGTTGAAGGATCTACCGGATTTGGGGGTGATACA 945
QY 287 SerIleGluGlyTyrIleIleGly-ProAlaThrGluAlaAlaAspLeuValThrAs 306
Db 946 TCTATTGAAGAGTACATATATGGACCTCTGACCAGGTGGTGTGATGGTGCATGAACATGC 1005

```



136 GATAATGCCACCATTCAGGATCTATTTTGTACATAAAGAGGAGTTTAAATTTACAAAGT 195  
 QY  
 96 GluProThrValGluGluLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
 Db  
 196 GAACCCACAATAGAGGCTCTAATCTGGCCATGTCACCTTATTGGAGCCACTGTCGTTACG 255  
 Db  
 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
 QY  
 256 ACATGCTCTGGACCCCTTATCAGACTGTGTAGCCGGCGTCTCTATGTTGATCATCTCTCT 315  
 Db  
 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrProAsnValTyrValLeuLeu 155  
 QY  
 316 ATCCCTTATTTGTTGTTCTCTGTCATGCTATGCTCTCCAAAGTTTATATCTCTCTC 375  
 Db  
 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
 QY  
 376 TTTGCAAGGCTTTTATGATGGATTAGGCATTTGGCGGTCACTTGGTACCTCTTTTAC 435  
 Db  
 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
 QY  
 436 ATATCCGACAGACGCTCCACCTGAGATTAGGGATTACTAAATACTCTTCCGACGTTCACT 495  
 Db  
 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
 QY  
 496 GGTTCCTGCTGGAATGTTCTTTCTTCTACTGTATGTTCTTTGCTATATCACTCACCAAGGCC 555  
 Db  
 216 ProAspThrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
 QY  
 556 CCNACTGGAGACTCATGTTAGTGTCTCTTCAATCCCTCTCTCATTTATTTGCACTC 615  
 Db  
 236 ThrIlePheTyrLeuProGluSerProArgTyrProLeuValSerIleGlyArgMetAlaGlu 255  
 QY  
 616 ACATATTTCTTGTGCTCGAATCTCAAGATGCTTGTGAGGCGATGTCAGNAAGCCGATGCTAG 675  
 Db  
 256 AlaLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
 QY  
 676 GCCAAGAAGGTTTGCACACCTGCGGCACAGAGATGTCGCTGGTGAGATGCTTTA 735  
 Db  
 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
 QY  
 736 CTAGTTGAGGCTCTTGAGTTGGGGTGATACCGTTATAGAGAGTACATAATTAGTCCA 795  
 Db  
 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
 QY  
 796 GCCAATGAATTCAGT-----GATGCAGACAGACATCAAAATTGTAT 834  
 Db  
 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
 QY  
 835 GGCACACGACAGAGTCAATCTGATTCGAACTGCCAACCTGTCACTGGACAAAGTTCCATTGGC 894  
 Db  
 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
 QY  
 895 -----CTTGATCTCGAAGGAGGAGCATGGCAATCAAAAGTGCT---CTAGTG 939  
 Db  
 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlySer 375  
 QY  
 940 GACCCTCTAGTAGAGTCTTTGGTAGTGTCCATGAGAAGCTCCCAAGAACAGGA----- 993  
 Db  
 376 MetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAla 395  
 QY  
 994 -----AGCACACTTTTCCACACTTTGGAAGCATGTTTAGTGTGGGGGAATCAACCT 1047  
 Db  
 396 LysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAsp 415  
 QY  
 1048 AGGAATGAAGATTGGGATGAAGAAAGCATTTGCCAGAGAGGAGATGATTATGTATCTGAT 1107  
 Db  
 416 GlyAlaGly--GlyAspTyrGluAsnLeuHisSerProLeuSerArgGlnAla 434  
 QY  
 1108 GCTGCTGACCCGACGATTCTGATGACAAATTTGCAGAGTCCATTGATCTCAGCTCAAGCA 1167  
 Db  
 435 ThrGlyAlaGluGlyAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454  
 QY  
 1168 AAGAGTCCGGA---AGGGACATGCTGCTCTGCCCCAGGGA-----AGCATG 1212  
 Db

455 ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474  
 QY  
 1213 AGGCAAGGTAGTCTTTTGT-----CAAGGAGAACTGCTGGAAACTCAGGAATGGTGGT 1266  
 Db  
 475 GlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494  
 QY  
 1267 CGATGGCAGCTAGCATGGAAATGGTCTCAAAACAGAAAGGTGT----- 1308  
 Db  
 495 GlyPheLysArgValTyrLeuHisGlnGluValProGlySerArgArgGlySerIle 514  
 QY  
 1309 ---TTCAAGAGAATAATTTCACCAAGAGGTGCTCTGATCTAGACGTGGGTCCATC 1365  
 Db  
 515 ValSerLeuProGly--GlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533  
 QY  
 1366 ATTTCACTTCAGGCTGTGATCGCCGACACTAACAGATGGTGAGATTGTCCAGGCTGCC 1425  
 Db  
 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
 QY  
 1426 GCTCTAGTAGTCAATCAGCCCTTTACCAAGAGGCTTATGCATCAGCAGCCAGTTGGA 1485  
 Db  
 554 AlaAlaMetValHisProSerGluValAlaLysLysGlySerArgTyrLysAspLeuPhe 573  
 QY  
 1486 CCAGCCATGATTCAATCCCAACAGCTGCAAAAGGCGCAGTTGGAGTGACCTTTTA 1545  
 Db  
 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593  
 QY  
 1546 GAGCCTGGAGTGAAGCATGCATGATTGTGGGGGTGGGGATTCAAATTTCTTCAACAGTTC 1605  
 Db  
 594 AlaGlyIleAsnGlyValLeuTyrTyrProGlnIleLeuGluGlnAlaGlyValAla 613  
 QY  
 1606 TCTGGTATAAATGGGTGCTCTACTATACACCTCAAAATTTCTTGAGCAGCAGGTTGGT 1665  
 Db  
 614 ValIleLeuSerLysPheGlyLeuSerAlaSerAlaSerIleLeuIleSerSerLeu 633  
 QY  
 1666 TATCTCTTTCAACCTGGGCTCTCTACTTCTGCTCTTTTCTTTTATCAGCAGCGTG 1725  
 Db  
 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653  
 QY  
 1726 ACACCTTGTGAATGTTACCTGTATAGCTGTGGCATGAGGCTCATGGATATATCTGCG 1785  
 Db  
 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
 QY  
 1786 AGAAGACCTTGTGCTGCTGACTCAATCCCGTCTTGATAGTATCACTTCTCATATTAGTA 1845  
 Db  
 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
 QY  
 1846 ATAGGAGAGTCTTGTGAATTCGACAGTACTATAAATGCTTTTATCTCACTTCGAGTGT 1905  
 Db  
 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713  
 QY  
 1906 ATTTGCTACTTCTGTTGTTTCGTTCATGGGCTTTGGGCCAAATTCCTTAATATCCTTTGTCA 1965  
 Db  
 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733  
 QY  
 1966 GAGATCTTCCCTACTCGAGTTCGTGCTATGCATAGCTATATATGCTTCTTACATTTTGG 2025  
 Db  
 734 IleGlyAsnIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753  
 QY  
 2026 ATATGTGCATCATCATGCTACCTACTCGCTCCAGTTATGCTTAATCTGTGGGCTTGGT 2085  
 Db  
 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
 QY  
 2086 GGTGCTCTTGTGATGATCGGTTGTGTCATCATAGCTTGGGTGTTTGTCTCTTAAA 2145  
 Db  
 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793  
 QY  
 2146 GTTCCAGAAACCAAGGCGCATGCCACTGGGAAGTTATCATTTGAGTCTTCTCTGTGGGCA 2205  
 Db  
 794 LysGlnAla 796  
 QY  
 2206 AAGCAAGCT 2214  
 Db

RESULT 14  
US-10-424-599-108435  
; Sequence 108435, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 108435  
; LENGTH: 2573  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_68931C.1  
US-10-424-599-108435

Alignment Scores:  
Pred. No.: 1.04e-235 Length: 2573  
Score: 2483.50 Matches: 505  
Percent Similarity: 76.69% Conservative: 84  
Best Local Similarity: 65.76% Mismatches: 121  
Query Match: 60.94% Indels: 58  
DB: 17 Gaps: 12

US-10-051-909-32 (1-800) x US-10-424-599-108435 (1-2573)

Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAenLeuLeuGlnGlyTrp	75
Db	37	ATGAGCGGAGCTGTTCTTGTGCTGCTGCTATCGCAACTTGTGCTGCAAGGATGG	96
Qy	76	AspAenAlaThrIle-----	80
Db	97	GATAATGCCACCATTCGAGGTCATATTTTATTCGGTTAGCCAAATTAAGATCA	156
Qy	81	-----AlaAlaValLeuTyrlleLysLysGluPhe	91
Db	157	CCAAGCTAATGGCTTTCTATGCTATGCAGGATCTATTTGTACATANAAGAGGAGTTT	216
Qy	92	GlnLeuGlnAenGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla	111
Db	217	AAATTACAAAGTGMACCACATAGAGAGGCTAATTTGTGGCCATGTCACCTATTGGAGCC	276
Qy	112	ThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeu	131
Db	277	ACTGTCGTACGACATGCTCTGGACCTTATCAGACTTGTAGGCGCGGCTCTATGTTG	336
Qy	132	IleLeuSerSerIleLeuTyrllePhePheSerGlyLeuIleMetLeuTrpSerProAenVal	151
Db	337	ATCATCTCTATACCTTTATTTTGTGGTCTCTGCTCATGCTATGCTCCAAATGTT	396
Qy	152	TyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeu	171
Db	397	TATATCTCTCTTTGCAAGGCTTTTATAGATGGATTAGGCATTTGTTGGCGGTACCTTG	456
Qy	172	ValProLeuTyrlleSerGluIleAlaProSerGluIleArgGlyLeuLeuAenThrLeu	191
Db	457	GTACCTCTTTACATATCCGAGACAGCTCCACCTCGAGATTAGGGGATTAATACTCTT	516
Qy	192	ProGlnPheSerGlySerGlyMetPheLeuSerTyrlleCysMetValPheGlyMetSer	211
Db	517	CCGAGATTCACTGTTCTGCTGGAATGTTCTTTTCTTACTGATGTTGCTTGTATGTCA	576
Qy	212	LeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe	231
Db	577	CTACCAAGGCCCCCAATTTGGAGACTCATGTTGGGTGTTCTTTCAATTCCTCTCTCAT	636

Qy	232	PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIleGly	251
Db	637	TATTTTGCACCTCAGCTATTCTTCTTGCCTGAATCTCCAAGATGGCTTTGTC--AAAGGC	693
Qy	252	ArgMetAlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspAspValSerGly	271
Db	694	CGGATGCTAGAGGCCCAAGAGTTTTCACACGACTTCGCGGCAGACAAGATGTCGCTGGT	753
Qy	272	GluLeuSerLeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyr	291
Db	754	GAGATGGCTTTACTAGTTGAGGCTCTTGGAGTTGGGCGTGATACCGCTATAGAAGAGTAC	813
Qy	292	IleIleGlyProAla-ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluG	311
Db	814	ATAATTGGTCCAGCTCAATGAATTCAGT-----GAGGCAGAGCA	852
Qy	311	nileThrLeuTyrlleGlyProGluGlyGlnSerTrpIleAlaArgProSerIleGlyPr	331
Db	853	GATCAAAATGTATGGGACAGCAGAGGTGTATCTCGATTGCTTAACCTGCTCACTGGACA	912
Qy	331	oileMetLeuGlySerValLeuSerLeuAlaSerArgHieGlySerMetValAenGlnSe	351
Db	913	AAGTTCATTGGT-----CTTGTATCTCGAGGGAAGCATGCAAAATCAAG	960
Qy	351	rValProLeuMetAspProIleValThrLeuPheGlySerValHieGluAenMetProGl	371
Db	961	T---GGTCTAGTGGACCTCTAGTGAACCTCTTTGGTAGTGTCCATGAGAAGCTCCCGA	1017
Qy	371	nAlaGlyGlySerMetArgSerThrLeuPheProAenPheGlySerMetPheSerValTh	391
Db	1018	AACAGGA-----AGCACCTTTTCCACACTTTGGGAAGCATGTTTAGTGTGG	1065
Qy	391	rAspGlnHieAlaLysAenGluGlnTrpAspGluGluAenLeuHieArgAspAspGluGl	411
Db	1066	GGGGAATCAACCTTAGGAATGAAGATTGGGATGAAGAAAGCATTCGCCAGAGGAGAGATGA	1125
Qy	411	uTyrlleAlaSerAspGlyAlaGly--GlyAspTyrlleGluAenLeuHieSerProLeuLe	430
Db	1126	TTATGATCTCTGATGCTGTTGACACCGACGATTCGTATGACAAATTTGCAGAGTCCATGAT	1185
Qy	430	uSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHieHieGlyHieArgGlySe	450
Db	1186	CTCAGCTCAAGCAACAAAGTCGGAA---AGGACATGCTCTGCTCTCCCGCAGGA-----	1237
Qy	450	rAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerTh	470
Db	1238	-----AGCATGAGCAAGTAGTCTTTTG-----CAAGGAGAACCTCTCGTGAAGATC	1284
Qy	470	rAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGluAenGl	490
Db	1285	AGGGAATTGGTGGTGGATGGCAGCTAGCATGGAATGCTCTGAACAAGAGGTGTT-----	1339
Qy	490	YArgLysGluGlyGlyPheLysArgValTyrlleGlnGlnGluGlyValProGlySerAr	510
Db	1340	-----TTCAAGAGAATATATTTTGACCAAGAGGTGCTCTGATATCTAG	1383
Qy	510	gArgGlySerIleValSerLeuProGly---GlyGlyAspValPheGluGlySerGluPh	529
Db	1384	ACGTGGGTCCATCATTTCACTTCCAGGCTGTGATGGCCGCACACTAACAGATGGTGAGAT	1443
Qy	529	eValHieAlaAlaLeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluPr	549
Db	1444	TGTCAGGCTGCGCTCTAGTGTAGTCAATCAGCCCTTTTACAACAAGGAGCTTATGCAATCA	1503
Qy	549	oArgMetSerAspAlaAlaMetValHieProSerGluValAlaAlaLysGlySerArgTr	569
Db	1504	GCAGCCAGTTGGACCGCATGATTCATTCATCCCAACAGCTGCAAAAGGCGGAGTTG	1563
Qy	569	pLysAspLeuPheGlyProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIl	589
Db	1564	GAGTGACCTTTTAGAGCTGGAGTGAAGCATGCTGATTTGTTGGGGTGGGATTCAAAT	1623
Qy	589	eLeuGlnGlnPheAlaGlyIleAenGlyValLeuTyrlleThrProGlnIleLeuGluGl	609



Db 1624 TCTTCAACAGTCTCTGGTATATAAATGGGGTGTCTTACTATACACCTCAAATTTCTTGAGCA 1683  
Qy 609 nAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLe 629  
Db 1684 GGCAGGTGTGGTATCTCTTCAAACTGGGCTCGGTCTACTTCTGGCTCTTTCT 1743  
Qy 629 uIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMe 649  
Db 1744 TATCAGCAGCGTGACAACTTGTGTATGTTACCTTGATAGCTGTGGCCATGAGGCTCAT 1803  
Qy 649 tAspLeuSerGlyArgPheIleLeuLeuGlyThrIleProIleLeuIleAlaSerLe 669  
Db 1804 GGATATATCTGGCAGAAAGACCTTGTCTGCTGACTCAATCCCGTCTGTAGTAGTATCAT 1863  
Qy 669 uValIleLeuValIleSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSe 689  
Db 1864 TCTCATATAGTAATAGGAGTCTTGTGGAATTGGACAGTACTATAAATGCTTTTATCTC 1923  
Qy 689 rThrValSerValIleValTyPheCysCysPheValMetGlyPheGlyProIleProAs 709  
Db 1924 AACTTCGAGTGTATTGTCTACTTCTTGTGTGTGGTCTATGAGTCTATATATGTC 1983  
Qy 709 nIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAl 729  
Db 1984 TATCCTTGTTCAGAGATCTTCCCTACTCGAGTTCGTGTCTATGATGCTATATATGTC 2043  
Qy 729 aPheThrPheThrIleGlyAspIleIleValThrTySerLeuProValMetLeuAsnAl 749  
Db 2044 TCTTACATTTTGTATATGTGACATCATGACCTACTGCTCTCCAGTTATGCTTAATTC 2103  
Qy 749 aIleGlyLeuAlaGlyValPheSerIleTyAlaValValCysLeuIleSerPheValPh 769  
Db 2104 TGTGGGCTTGTGGTGTCTTGTGTATATGCGGTGTGTGTCATCATAGCTTGGGTGT 2163  
Qy 769 eValPheLeuLysValProGluThrTyGlyMetProLeuGluValIleThrGluPhePh 789  
Db 2164 TGTCTTCTTAAAGTTCAGAAACCAAGGCGCATGCCACTGGAAGTTATCATTGAGTCTT 2223  
Qy 789 eAlaValGlyAlaLysGlnAla 796  
Db 2224 CTCTGTGGGGCAAGCAAGCT 2245  
RESULT 15  
US-10-425-114-29146  
; Sequence 29146, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 29146  
; LENGTH: 2271  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB83-005-C10\_FLI  
US-10-425-114-29146  
Alignment Scores:  
Pred. No.: 1,01e-231 Length: 2271  
Score: 2442.50 Matches: 472  
Percent Similarity: 83.66% Conservative: 76  
Best Local Similarity: 72.06% Mismatches: 100

Query Match: 59.94% Indels: 7  
DB: 17 Gaps: 6  
US-10-051-909-32 (1-800) x US-10-425-114-29146 (1-2271)  
Qy 143 LeuIleMetLeuTrpSerProAsnValTyValLeuLeuAlaAatqPheValAspGly 162  
Db 3 CTGTGTATGCTCTGGGCGCAAGTGTACATCTTCTCTCTCGCAGGCTCATGTATGGG 62  
Qy 163 PheGlyIleGlyLeuAlaValThrLeuValProLeuTyIleSerGluIleAlaProSer 182  
Db 63 TTTGCGTATCGGTGTGGCGTCACTTGTTCCTCTCTATCATCTCCGAGCTGCGCGGACA 122  
Qy 183 GluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeu 202  
Db 123 GACATTCGTGGGTGTGTAACACGTTGCGCAGTTCAGTGGGTGAGGAGGATGTTCCTC 182  
Qy 203 SerTyCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeu 222  
Db 183 TCTACTGTGATGGTGTGGGATGTCCTCATGCCAACCTGATTGGAGGCTCATGCTT 242  
Qy 223 GlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyIleProGlu 242  
Db 243 GGAGTCTGTGCGATCCCGTCACTTATTACTTTGGACTGACTGTCTTCTACTTGCCTGAA 302  
Qy 243 SerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysValLeuGlnLys 262  
Db 303 TCACCAAGGTGGCTTGTGAGCAAGAGAGATGGCTGAGCGCAAGAGAGTGTTCGAAAG 362  
Qy 263 LeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGluVal 282  
Db 363 CTGCGGGAAGAGAAGATGTCTCAGGGAGATGGCTCTTCTAGTTGAAGGTTTGGGGTC 422  
Qy 283 GlyGlyAspThrSerIleGluGluTyIleIleGlyProAlaThrGluAlaAlaAsp 302  
Db 423 GGTAAGATACACGATTTGAAGAATAACATAATTTGGTCCCGATGATGAATCTGCTGATGAA 482  
Qy 303 LeuValThrAspGlyAspLysGluGlnIleThrLeuTyGlyProGluGluGlnSer 322  
Db 483 GGGTGGCTCCA--GATCCAGAGAAGATCAAACTATATGGACCTGAAGAAGCCCTATCT 539  
Qy 323 TrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer 342  
Db 540 TGGGTTCGCCGACCTGTTGGGCAAGAGTGTCTTTGGAAGCGGCTTAGTCTCATCTCT 599  
Qy 343 ArgHisGlySerMet---ValAsnGlnSerValProLeuMetAspProIleValThrLeu 361  
Db 600 CGTCATGGGAGTATGGCGGCTAGTCAGGGTAAAGCCCTCTGTGGATCTATGGTCACTCT 659  
Qy 362 PheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhe 381  
Db 660 TTCGGAAGTGTTCATGAAAGATGCCCTGAGATCATGCGGAGCATGAGAGACATTTGTTT 719  
Qy 382 ProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluInTrpAsp 401  
Db 720 CCCAACTTTGGCAGCATGTTTAGTGTTCGCGCACGAGCAGCAGAGCTGAGCTGGGAC 779  
Qy 402 GluGluAsnLeuHisArgAspAspGluGluTyIleAlaSerAspGlyAlaGlyGlyAspTyr 421  
Db 780 GCCGAGAGT---CAAAGGGAAGGTGAAGATTATCTTCGGATCATGTCGCGCATGACATT 836  
Qy 422 GluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAsp 441  
Db 837 GAGGATAACCTCCAAAGCCCACTTATTCGTCTCAGGCCAACAAAGTGTGAAGGAAGAG 896  
Qy 442 IleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGly 461  
Db 897 ATCGCT-----GCACCTCATGTTAGCATATTTGGTGTCTGGGAAGGAGCAGTAGCTTG 950  
Qy 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTyTrpGlnLeuAlaTrpLys 481  
Db 951 CAGGGTGGGAGGCGCAGTAAAGCAGCATGGGCAATTGCGGAGGATGGCAGTTGGCGTGGAAA 1010





This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 12:44:19 ; Search time 5523.13 Seconds

(without alignments)  
5513.431 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSLAVQTFTPDLR.....PLEVITEFPVAGKQAANA 800

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cpn2.1/USPTO spool\_p/US10051909/runat\_13042005\_074037\_14060/app\_query.fasta\_1.1678  
-DB=EST -QPMF=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdd -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN 1.1 6628 @runat\_13042005\_074037\_14060 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hcc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse1:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3027	74.3	2180	3 AY105508	AY105508 Zea mays
2	2364	58.0	2292	9 CL980871	CL980871 OsIFCC045
3	1934.5	47.5	1932	9 CL968379	CL968379 OsIFCC040
C 4	1399	34.3	911	8 CC423708	CC423708 PUH0R49TB
5	1382.5	33.9	845	8 BZ723844	BZ723844 PUCFI60TD
C 6	1279	31.4	795	8 BZ989776	BZ989776 PUGDY07TD
7	1268.5	31.1	1135	3 CENS0A3H4	BX828912 Arabidops
8	1168	28.7	833	6 CA065405	CA065405 SCACAD103
9	1154	28.3	733	6 CA146057	CA146057 SCVPR207

10	1142	28.0	696	6 CA075335	CA075335 SCJFAM106
11	1120	27.5	1082	7 BQ208781	BQ208781 FGAS02050
12	1101	27.0	829	5 BQ752737	BQ752737 WHE4118_E
13	1100	27.0	871	6 CA170109	CA170109 SCQGB108
14	1086	26.7	1107	5 BX841312	BX841312 BX841312
15	1066.5	26.2	848	7 CF445664	CF445664 EST682009
16	1063.5	26.1	769	6 CA243478	CA243478 SCQFL407
17	1055	25.9	994	6 CA067439	CA067439 SCQAD106
18	1034	25.4	636	6 CF018228	CF018228 QBM9003.X
19	1033.5	25.4	950	7 CK263963	CK263963 EST710041
20	1030	25.3	692	6 CA112602	CA112602 SCEQLB106
21	1022	25.1	631	6 CA252785	CA252785 SCBFFL114
22	1018	25.0	787	7 CF182124	CF182124 JG-2 Chin
23	1014.5	24.9	643	6 CA263039	CA263039 SCPRUB202
24	1014.5	24.9	858	6 CA207025	CA207025 SCQGB1C0
25	1009	24.8	644	6 CA279697	CA279697 SCCCLF800
26	1003	24.6	616	6 CA081000	CA081000 SCCCAM200
27	997	24.5	606	6 CA230633	CA230633 SCJFFL3C0
28	990.5	24.3	745	6 CB635485	CB635485 OSI1Bb150
C 29	981	24.1	594	1 AI861088	AI861088 603011H02
30	980	24.0	620	6 CA167255	CA167255 SCCGB100
31	966	23.7	669	6 CA073177	CA073177 SCEPAM105
32	957	23.5	817	8 BZ819234	BZ819234 PUGDY07TB
33	945	23.2	903	7 CF211894	CF211894 CGF100066
34	941.5	23.1	655	6 CA268224	CA268224 SCQSR305
35	939	23.0	701	6 CD879843	CD879843 AZ04.1061
36	937	23.0	827	6 CA207188	CA207188 SCEQSB1C0
37	935	22.9	801	7 CK201790	CK201790 FGAS01031
38	930	22.8	859	4 BM817360	BM817360 HC106B05
39	928	22.8	749	5 BU040051	BU040051 PP L6a000
40	924	22.7	688	6 CA502649	CA502649 WHE4338_A
41	921.5	22.6	700	5 BQ870435	BQ870435 QGD9A19.Y
42	921.5	22.6	703	6 CA502490	CA502490 WHE4048_B
43	921	22.6	687	6 CD893252	CD893252 G118.123E
C 44	916	22.5	764	7 CO983244	CO983244 GM89018A2
45	916	22.5	847	6 CB981615	CB981615 CAB70005_

ALIGNMENTS

RESULT 1					
LOCUS	AY105508	2180 bp	mRNA	linear	HTC 16-OCT-2002
DEFINITION	Zea mays PC0114533 mRNA sequence.				
ACCESSION	AY105508				
VERSION	AY105508.1 Gi:21208586				
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.				
AUTHORS	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
TITLE	Unpublished (2002)				
JOURNAL	2 (bases 1 to 2180)				
REFERENCE	Coe,E.H.				
AUTHORS	Direct Submission				
TITLE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
JOURNAL	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizenap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.				
FEATURES	Location/Qualifiers				
source	1..2180				
	/organism="Zea mays"				

```

/mol_type="mRNA"
/db_xref="MaizeDB:636667"
/db_xref="taxon:4577"
/cloned_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1,066-294 Length: 2180
Score: 3027.00 Matches: 591
Percent Similarity: 99.66% Conservative: 2
Best Local Similarity: 99.33% Mismatches: 2
Query Match: 74.28% Indels: 0
DB: 3 Gaps: 0

US-10-051-909-32 (1-800) x AV105508 (1-2180)
QY 206 MetValPheGlyMetSerLeuSerProSerProAspTTPAArgileMetLeuGlyValLeu 225
DB 50 ATGGGTGTTGGGATGTCCTCTGCGCATCACCCGATTCGAGAAATATGCTTGGTGTGCTC 109
QY 226 AlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArg 245
DB 110 GCGATACCTTCATGTTCTCTTGGTTGCAATATTTATCTCTCCGAAATCTCCAAGA 169
QY 246 TrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGly 265
DB 170 TGGCTCGTTAGCAAAAGGTGGATGGCAGAGGCAAAAAGGTGTTGCCAAAAGTTACGGTGG 229
QY 266 LysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAsp 285
DB 230 AAAGACGATGCTCAGGTGAATTTGTCCTTCTTCGAAAGGGTTGGAGTTGGAGGAGAC 289
QY 286 ThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThr 305
DB 290 ACTTCCATTGAAGAGTACATCATTCGACCTGCCACCGAGGCGAGCGCATGATCTTTGTA 349
QY 306 AspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAla 325
DB 350 GACGGTGATAAGAACAAATCACATTTATGGGCTCGAAGAGCCAGTCATGCAATTGCT 409
QY 326 ArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGly 345
DB 410 CGACCTTCTAAGGGACCCATCATGCTTGAAGTGTGCTTTCTCTGCAATCTCGTCATGGG 469
QY 346 SerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerVal 365
DB 470 AGCATGGTGAACACAGAGTGATACCCCTTATGGATCCGATTTGACACTTTTGGTAGTGC 529
QY 366 HisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGly 385
DB 530 CATGAAATATGCCTCAAGCTGGAGAAAGTATGAGGAGCAATTTGTTTCCAAACCTTTGGA 589
QY 386 SerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeu 405
DB 590 AGTATGTTTCAGTGTCCACATCAGCATGCCAAAATGAGCAGTGGGATGAAGAGAACTCT 649
QY 406 HisArgAspAsnGluGlnTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeu 425
DB 650 CATAGGGATGACGAGGAGTACGCACTGATGGTCAGGAGGTGACTATGAGGACAACTCT 709
QY 426 HisSerProLeuLeuSerArgGlnIleThrGlyAlaGluGlyGlyAspIleValHisHis 445
DB 710 CATAGCCCAATGCTGTCCAGGACGAGCAGGCTGGGAGGAGGAGCAATTTGCAACCAT 769
QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAsp 465

```

```

DB 770 GGTACCCGTGGAGGTGCTTTTGAGCATGAGAAAGGCAAAAGCCCTTTAGGGAGGGTGGAGAT 829
QY 466 GlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLys 485
DB 830 GGTGTGAGCAGCACTGATATCGGTGGGGATGGCAGCTTGTCTGGAAATGTCAGAGAAG 889
QY 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
DB 890 GAAAGTGAATGGTAGAAGAAAGGTGGTTTCAAAAGAGTCTACTTTGCCAAAGAGGGA 949
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525
DB 950 GTTCTGGCTCAAGAAAGGGCTCAATTTGTTCACTTCCCGGTGGTGGGATGTTCTTGAG 1009
QY 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
DB 1010 GGTAGTAGTGTGTACATGCTGCTGTCTTAGTAGTCAGTCAGCACCTTTCTCAAGAGGT 1069
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
DB 1070 CTTGTCTGAACCAACGATGCTGCCATGCTTCAACCATCTGAGGTAGCTGCCAAA 1129
QY 566 GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal 585
DB 1130 GGTTCACGTTGGAAAGATTTGTTGAACCTGGAGTGGAGCGTGCCTGTAGTCGGTGT 1189
QY 586 GlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrThrProGln 605
DB 1190 GGAATTCAGATCTTCAACAGTTGCTGGAATAAACGGTGTCTGTACTATACCCACAA 1249
QY 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
DB 1250 ATCTTGAGCAAGCTGGTGGCAGTTATTTCTTCCAAATTTGGTCTCAGCTCGGCATCA 1309
QY 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
DB 1310 GCATCCATCTTGATCAGTCTCTCACTACCTTACTATGCTTCTTGCAITGGCTTTGCC 1369
QY 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 665
DB 1370 ATGCTGCTTATGATGATCTTTCGGAAGAAAGGTTTGTGCTAGGCACAAATTCGAATCTG 1429
QY 666 IleAlaSerLeuValIleLeuValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
DB 1430 ATAGCATCTCTAGTTATCTGCTGTCCTAATCTAATTTGGGTACACTAGCCCAT 1489
QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
DB 1490 GCTTTGCTCTCCACCATCAGTGTATCTGCTACTTCTGCTCTTCGTTATGGGATTTGGT 1549
QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
DB 1550 CCATCCCAACATTTTATGTCAGAGATCTTCCACACAGGTTCTGTGGCTCTGTATT 1609
QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745
DB 1610 GCATTTGTGCTTTTACATCTCGATCGAGATATCATCGTACCTACAGCTTCTGTGTG 1669
QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
DB 1670 ATGCTGAATGCTATTGGATCGCGGGTGTTCAGCATATATGCACTGATGCTGTATT 1729
QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
DB 1730 TCTTTGTGCTTCTTCTTAAAGGTCTCGAGACAAAGGGGATGCCCTTTGAGGTATT 1789
QY 786 ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
DB 1790 ACCGAATCTTTTGCAATTTGGTGGAGCAAGCGGCTGCCAAAGCC 1834

```

RESULT 2  
CL980871  
LOCUS

CL980871 2292 bp DNA linear GSS 21-SEP-2004

**DEFINITION** OsIFCC045538 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.  
**ACCESSION** CL980871  
**VERSION** CL980871.1 GI:52416230  
**KEYWORDS** GSS.  
**SOURCE** Oryza sativa (indica cultivar-group)  
**ORGANISM** Oryza sativa (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
**REFERENCE** 1 (bases 1 to 2292)  
**AUTHORS** Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.  
**TITLE** An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis  
**JOURNAL** Unpublished (2004)  
**COMMENT** Contact: Chen Chen  
 Department of Bioinformatic  
 Beijing Institute of Genomics  
 Chinese Academy of Sciences, Beijing 101300, China  
 Tel: 86-10-80481559  
 Fax: 86-10-80488676  
 Email: chenchen@genomics.org.cn  
 Rice genomic sequence.  
 Class: exon-trapped.  
**FEATURES**  
 source  
 1..2292  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39946"  
 /clone\_lib="Oryza sativa Express Library"  
 /notes="Oryza sativa exon trapped genomic sequences"

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 1,54e-227 Length: 2292  
 Score: 2364.00 Matches: 482  
 Percent Similarity: 75.13% Conservative: 95  
 Best Local Similarity: 62.76% Mismatches: 147  
 Query Match: 58.01% Indels: 44  
 DB: 9 Gaps: 16  
 US-10-051-909-32 (1-800) x CL980871 (1-2292)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75  
 DB 1 ATGAGAGGCCCGGTGTGTGGTGTGGTGTGGCGCCGCCATCGCACTACCTGCGAGGATGG 60  
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleIleLysLysGluPheGlnLeuGlnAsn 95  
 DB 61 GACACGCCACCATCGCGCGCGGTGCTGTACATCAAGCGGAGTTGGCCCTCGAGACC 120  
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
 DB 121 CAGCCCGCGCTGGAGGCGCTCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 180  
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
 DB 181 ACCTTCTGTGGCCCGGTCTCGACCTGTGTGGCGCGCGCCGCCATGCTCATCGGCTCTCTCC 240  
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
 DB 241 CTCTCTACTTCGCGCGCGGCTCATCATGCTCTGTGTCTCCCAACGCTACGTCCTCTCC 300  
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
 DB 301 CTGCGCGCGCTGTGACGCGCTTCGGCGTTCGGCGTTCGGCGTTCGGCGTTCGGCGTTCAC 360  
 QY 176 IleSerGluIleAlaProSerGluIleAlaArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
 DB 361 ATCTCCGAGACCTCCCGCGCGGATCCCGCGCGGCTCAACACACTGCGCGCATTCACC 420

QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
 DB 421 GGCTCCGCGCGCATGTTCACTACTGATGATGATGATGATGATGATGATGATGATGATGATG 480  
 QY 216 ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
 DB 481 CCCNACTGGGCATCATGCTCGCGTCTCTTCGTCCTCGCTCGCTCGCTCGCTCGCTCGCTC 540  
 QY 236 ThrIlePheTyrLeuProGluSerProArgTyrTrpLeuValSerLysGlyArgMetAlaGlu 255  
 DB 541 ACCGTCTTCTACCTCCCGAGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
 DB 601 GCCAGGTTGTGCTCGAGATGTCGCGCGCGCGAGAGCTCTCCCGCGAGATGAGCGCTG 660  
 QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTyrIleIleGlyPro 295  
 DB 661 CTCGTGAGGGCTCGGCACCGCGCGGCGACCCGAGATCGAGGACTACGTCGTGCGGCCG 720  
 QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGln----- 311  
 DB 721 TCCGAGGCGGACGCC-----GGCGAGAACGAGCAGCGAGGAGGACACG 762  
 QY 312 IleThrLeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGly--- 330  
 DB 763 GTACGCTGTACGCGCGGAGCGGCTTGTGGTGGCGCGCGGCTGCGCGCGCGCG 822  
 QY 331 ---ProIleMetLeuGlySerValLeuSerLeu---AlaSerArgHisGlySerMetVal 348  
 DB 823 CGCGGCGAGCATGCTGGGAGCTCGCTGGGGTGCAGCGCTGCGGCCACCGCAGCATGTAC 882  
 QY 349 AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368  
 DB 883 GAGCAG-----ATGAAGGACCCCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 933  
 QY 369 MetProGlnAlaGly-----GlySerMetArg---SerThrLeuPheProAsn 383  
 DB 934 CTGCGCGAGTCCGCGCGCGCGCCACCGCGAGCATGAGGGGAGCAGCTGTGTTCCCAAC 993  
 QY 384 PheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGlu 403  
 DB 994 CTCGGAGGATGCTTAGCGTCAACGACAGGCCCGCG---GGCAGCAGCTGGGACGAGGAG 1050  
 QY 404 Asn-----LeuHisArgAspAspGluGlyTyrAlaSer---Asp 415  
 DB 1051 AACGTGCGAGCTCGCGCAGCAGCCTCGAGGAGGAGGAGGAGTACCTCTCCGACGAC 1110  
 QY 416 GlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThr 435  
 DB 1111 GGCAAGGACGACGACGACGCGCGCGCTGCGAGCGCGCTGTGTCGCGCGAGAGCACC 1170  
 QY 436 GlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSerMetArg 455  
 DB 1171 GAGCTGGAACCAAGAACGAGCCCGCTCGGCCAG-----GTGGCGATGACG 1218  
 QY 456 ArgGlnThrLeuLeuGlyGluGlyGly-----AspGlyValSerSerThrAspIleGly 473  
 DB 1219 CGGCACGACGACATCGCGCGCGCGCGCGCTCGAGACGCGCAGCACCATGGGCATCGGC 1278  
 QY 474 GlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493  
 DB 1279 GCGCGGTGGCAGCTGCGTGGAAAGTGGACGAGAACGTGGGCGCCCGCAGCGCTCAAGCGC 1338  
 QY 494 GlyGlyPheLysArgValTyrLeuHisGln-----GluGlyValProGlySerArgArg 511  
 DB 1339 GCGCGGTGAAGCGCATGATACCTGACAGAGAGTCCGAGCCCGCGCGCGCGCGCGCGCG 1398  
 QY 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGlyGlySerGluPheValHis 531  
 DB 1399 GCG 1455  
 QY 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly---LeuAlaGluProArg 550

```

Db      1456  GCGGGGGCGCTGGTGGAGCGGTGCGATGTGTACCAAGGACGCTCTGATCGGCGAGAGC 1515
Qy      551  MetSerAspAlaAaMetValHisProSerGlu-----ValAlaAlaLysGly 566
Db      1516  CCGACCGAGCGCGCTTCGCGAACCGCGCGAGCGGTGGCGGGCGGTGCGACGGGC 1575
Qy      567  SerArgTTPyAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly 586
Db      1576  CCGCGCTGGCGGAGCTGCTCGAGCGGGCGCTCCGCCACGCGCTCTCTCGGGGTAC 1635
Qy      587  IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrProGlnIle 606
Db      1636  ATCCAGATCTCCACAACTTCGCGCAATCCCGGCAACCGCGGTCTCTACTACCCCGCAGATC 1695
Qy      607  LeuGluGlnAlaGlyValAlaValIleLeuSerIysPheGlyLeuSerSerAlaSerAla 626
Db      1696  CTCGACCAAGCGCGGTGAGCGTCTCTCCGCCAGCTCGGCCCTCTCCGGCGACTCCACC 1755
Qy      627  SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db      1756  TCCATCTCTCATCAGCGGCTCACCACGCTCTCATGCTCCGCTCCATCGGCGTCCGCA 1815
Qy      647  LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db      1816  CGCCTCATGAGCGCTCGGGCGCGCGCGCTCTCTCTGAGCGCTGCGCGTCTCTGTC 1875
Qy      667  AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db      1876  GGTCTCTCGCGTCTCTGTGTGGGAACTGCTGTCGTCATGGCGGCGCGCGGCGCG 1935
Qy      687  LeuLeuSerThrValSerValIleValTyrPheCysPheValMetGlyPheGlyPro 706
Db      1936  GGCCTGTGACGGGAGCGTATGCTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1995
Qy      707  IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
Db      1996  ATCCCAACATCTCTCGCGCGAGATCTCCCGACGAGGGGTGAGGGACTCTGCTGTC 2055
Qy      727  IleCysAlaPheThrPheThrIleGlyAspIleIleValThrTyrSerLeuProValMet 746
Db      2056  ATCTGCTCGCTGACGTTCTGCTGCGGCGACATCGCGCTCAAGTACAGCCCTCCCGCT 2115
Qy      747  LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766
Db      2116  CTCAGCTCGTGGGCTCGCGGGGTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2175
Qy      767  PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db      2176  CTCGTTCTCGTGGGCTCAAGTGTCCCGAGACCAAGGGCTCCCGCTCGAGGTCATCATC 2235
Qy      787  GluPhePheAlaValGlyAlaLys 794
Db      2236  GAGTCTTCAACGTCGGGCGCCAG 2259

RESULT 3
LOCUS      CL968379
DEFINITION OsIPCC040330 Oryza sativa Express Library Oryza sativa (indica
ACCESSION  CL968379
VERSION     CL968379.1 GI:52391393
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
            1 (bases 1 to 1932)
REFERENCE   1
AUTHORS     Ma L., Wang J., Chen C., Liu X., Su N., Li L., Wang X., Cao M.,
            Jiao Y., Sun N., Zhang X., Bao J., Sun D., Zhao H., Yuan L.,
            Wong G. K. S., Deng X. W. and Wang J.
TITLE       An analysis of transcriptional regulation of the rice genome and

```

its comparison to Arabidopsis  
Unpublished (2004)  
Contact: Chen Chen  
Department of Bioinformatic  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.

JOURNAL  
COMMENT

FEATURES  
source

Location/Qualifiers  
1..1932  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:39946"  
/clone\_lib="Oryza sativa Exress Library"  
/note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores:  
Pred. No.: 3.59e-184 Length: 1932  
Score: 1934.50 Matches: 403  
Percent Similarity: 67.07% Conservative: 94  
Best Local Similarity: 54.39% Mismatches: 131  
Query Match: 47.47% Indels: 113  
DB: 9 Gaps: 11

US-10-051-909-32 (1-800) x CL968379 (1-1932)

```

Qy      56  MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db      1  ATGAGAGCGCGGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Qy      76  AsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPhe---GlnLeuGln 94
Db      61  GACACGCGCCATCTCCCGCGCTCTCTCTACATGAGGAGGACTTCGCGCGCTCCAA 120
Qy      95  AsnGluProThrValGluGlyLeuIleValSerMetSerIleGlyAlaThrIleVal 114
Db      121  GCCACCCCGACATCCAGGGTCTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy      115  ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134
Db      181  ACCACCTTCTCCGCGCGCTCTCCGACTCCCGAGCGCGCGCGCGCGCGCGCGCGCT 240
Qy      135  SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeu 154
Db      241  GCGTGTCTTACTCCTCTCGCGGCTCTCTATGCTATGCTATGCTATGCTATGCTATG 300
Qy      155  LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db      301  CTACTCGCGCGCTTGTGACGGCTTCGCCATCGCGCTTCGCCGCTCGCTCGCTCGCT 360
Qy      175  TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db      361  TACATCTCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 420
Qy      195  SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db      421  ACCGCTCCACCGGATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Qy      215  SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
Db      481  ATCCCCAACTCGCGCTCATGCTCGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Qy      235  LeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerIysGlyArgMetAla 254
Db      541  CTCACCATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Qy      255  GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSer 274
Db      601  GAGGCCAGGACCGTCTCCAGATGCTCAGGGGACGCCAGGACGTCGCGCGAGATGGCG 660

```

```

QY 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGly 294
DB 661 CTCCTCTGTCAGGGGTCACACCGGACGCGACACGCGCATCGAGGAGTACGTCGCGG 720
QY 295 ProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeu 314
DB 721 CCACCGACGAGGCC----- 735
QY 315 TyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeu 334
DB 736 -----TCGGATTGGCCCGCAGGCAGC-----ATGTTT 762
QY 335 GlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeu 354
DB 763 GGCAGCGCGCTCAGCTGGCATCGCACAGGCGAGCATGCTGCACCAC-----CTC 813
QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
DB 814 AAGGACCCCGTGTCCGCTCTCTGACAGCCTCCACGACATGAAACCCCGCCGCGTGGC 873
QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
DB 874 -----ACTACTGACGTCCCAACCTCGGCAGCATGATCGCGCTCCACGACAGCGCG 924
QY 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
DB 925 CCCATCGAC-----TGGGACGAGGAGAAC----- 948
QY 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
DB 949 -----TCGGGGACGACGACGTCGATCGCCGCGCGCTGTAAACCATG----- 993
QY 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
DB 993 ----- 993
QY 455 ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
DB 994 -----GAGGGCGAGCGCGCGACGACGCGTGGCGCATCGCGCGG 1032
QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494
DB 1033 GGGTGGCAGCTGCGTGGAGTGGACGAGGGCGCTCGCGCGGACGCGGACGAGCAGC 1092
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
DB 1093 ACGGTCAACAGGATGTACCTGCACGAGGAGCAGCGCGG----- 1134
QY 515 ValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla 534
DB 1135 -----GTGCACGCGCGCGG 1149
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
DB 1150 CTGGTGAGCAGTCGCGCTGTGTGCACCAAGAGGCGGAGCGGAGGTGGAGGTGGG 1209
QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574
DB 1210 -----TGGCGGAGGTGTGGAG 1227
QY 575 Pro---GlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
DB 1228 CCGCGCGGCTCGGCACGCGTGTGTGCGGCGTGCAGATCCGATCCGACGAGTTC 1287
QY 594 AlaGlyLeuAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
DB 1288 TCGGGCATCAGCGCGTGTCTTGTACACCGCCAGATATGAGCAGCGCGCGCGTGGG 1347
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
DB 1348 GTGCTGCTCTCCAGGGTGGCGCTCGCGGAGCTCCGCTCCATCTCATCAGCGCGT 1407

```

```

QY 634 ThrThrLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
DB 1408 ACACGCTGCTGATGTGCCACGATCGCGGTGGCCATCGCGCTCATGGACGTGTCGGG 1467
QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
DB 1468 CGCGGAGGCTGCTCTGTGGACGATCCCGCTGCTGCTGGCGTCTGGCGGTGCTGGT 1527
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
DB 1528 GCGCGGAGGTGGCGCGATGCGCGCGCGCGCGACGCGCGTGTGCACGGGAGCGTG 1587
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
DB 1588 GTGGTGATCTGCTGCTGCTGCTGATGATGCGGCTTCGCATCGCCATCTGCTCGCTTCTGG 1647
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
DB 1648 GAGATATTCACACGCGCGTGGTGGCTTCTGCATCGCCATCTGCTCGCTCGCTTCTGG 1707
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
DB 1708 CTCGCGACATCGCGCTACCTACACCTCCCGCTCATGCTCGCTCCCTCGCGCTCGCC 1767
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
DB 1768 GGCCTCTTCGCATCTACCGCGCGCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCGG 1827
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
DB 1828 GTGCCGAGACCAAGGCGCTCCCGCTCGAGGTGCATCATCGACTTCTTCAACGTGCGAGCC 1887
QY 794 Lys 794
DB 1888 AAA 1890

RESULT 4
CC423708/c
LOCUS
DEFINITION
CC423708
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 911)
Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUH049TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
source
1..911
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM0610.0"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high

```

CoT selected genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 1.78e-130 Length: 911  
Score: 1399.00 Matches: 267  
Percent Similarity: 98.90% Conservative: 3  
Best Local Similarity: 99.16% Mismatches: 2  
Query Match: 34.33% Indels: 0  
DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x CC423708 (1-911)

270 SerGlyGluLeuSerLeuLeuGluGlyLeuValGlyAspThrSerIleGlu 289  
QY ::::: 267  
Db ACAGGTGAATTCCTCTTCGAAAGGTTGGAGTTGAGGAGACATTCATTGAA 759  
290 GluTyrIleGlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLys 309  
QY ::::: 267  
Db GAGTACATCAATGGACCTGCCACGAGGAGCCGATGATCTTTGCTGACGGTGATAAG 699  
310 GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329  
QY ::::: 267  
Db GAACAAATCACACTTATGGGCTGAAGAAGCCAGTCATGGATTGCTCGACCTTCCAA 639  
330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349  
QY ::::: 267  
Db GGACCCAGCATCTTGGAGGTGCTTCTTCTGATCTCTGATGAGGAGCATGGTGAAC 579  
350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369  
QY ::::: 267  
Db CAGAGTGTACCCCTTATGGATCCGATTTGTGACATCTTTTGGTAGTGTCCATGAGAATATG 519  
370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389  
QY ::::: 267  
Db CCTCAAGCTGGAGGAATATGAGGAGCACATTTGTTCCAAACTTTGGAAGTATGTTCCAGT 459  
390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp 409  
QY ::::: 267  
Db GTCCAGGATCAGCATGCCAAATGAGGAGTGGGATGAAGAATCTTTCATAGGAGTAC 399  
410 GluGluTyrAlaSerAspGlyValGlyGlyAspTyrGluAspLeuHisSerProLeu 429  
QY ::::: 267  
Db GAGGAGTACGATCTGATGTTGTCAGAGAGTGCTATGAGGACATCTCCATAGGCCCATG 339  
430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly 449  
QY ::::: 267  
Db CTGTCCAGGAGGACACAGTCCGAGGAGGAGGACATTTGTCACCATGTCACCGTGA 279  
450 SerAlaLeuSerMetArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSer 469  
QY ::::: 267  
Db AGTGCTTTTGTAGCATGAGAGGCAAGCCTCTTATGGGAGGTTGAGATGGTGGAGCAGC 219  
470 ThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsn 489  
QY ::::: 267  
Db ACTGATATCGTGGGGATGCGCAGCTTCTTGGAAATGCTCAGAGAAGGAGTGGAGAAT 159  
490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509  
QY ::::: 267  
Db GGTAGAAAGGAAGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTGGCTCA 99  
510 ArgArgGlySerIleValSerLeuProGlyGlyAspValPheGluGlySerGluPhe 529  
QY ::::: 267  
Db AGAAGGGGCTCAATGTTTTCATCTCCGGTGGTGGCGATGTTCTTGAGGGTAGTGATTT 39  
530 ValHisAlaAlaLeuValSerGlnSerAlaLeu 541  
QY ::::: 267  
Db GTACATGCTGCTGCTTTTAGTAAGTCAGTCAGCACATT 3

RESULT 5

BZ723844

LOCUS

DEFINITION

BZ723844 845 bp DNA linear GSS 24-FEB-2003  
PUCF160TD ZM\_0.6\_1.0\_KB Zea mays genomic clone ZMMBta133J23,

genomic survey sequence.

ACCESSION BZ723844  
VERSION BZ723844.1 GI:28519500  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 845)  
White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and  
Bennetzen, J.

TITLE Maize Genomics Consortium  
JOURNAL Unpublished (2003)  
COMMENT Contact: Cathy Whitelaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TP  
Class: sheared ends.

FEATURES

Location/Qualifiers  
1..845

/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBta133J23"  
/clone\_lib="ZM\_0.6\_1.0\_KB"  
/note="Vector: pCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
CoT selected genomic DNA library"

ORIGIN

Alignment Scores:  
Pred. No.: 7.4e-129 Length: 845  
Score: 1382.50 Matches: 274  
Percent Similarity: 97.52% Conservative: 1  
Best Local Similarity: 97.16% Mismatches: 6  
Query Match: 33.93% Indels: 3  
DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x BZ723844 (1-845)

303 LeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlnSer 322  
QY ::::: 274  
Db 3 CTTGTTACTGAGCGTGATAAGGAACAAATCACATTTATGGGCTGAAGAAGGCCAGTCA 62  
323 TrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer 342  
QY ::::: 274  
Db 63 TGGATTGCTCGACCTTCCAGGGACCCAGCATGCTTGGAGTGTCTTCTTGCATCT 122  
343 ArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPhe 362  
QY ::::: 274  
Db 123 CGTCTAGGAGCATGGTGAACAGAGTGTACCCCTTATGGATCCGATTTGACACTTTT 182  
363 GlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePro 382  
QY ::::: 274  
Db 183 GGTAGTGTCCATGAGAATATGCTCAAGCTGGAGAGATGATGAGGAGCACATTTTCCA 242  
383 AnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGlu 402  
QY ::::: 274  
Db 243 AACTTTGGAAGTATGTTTCAGTGTACAGATGCTGTAATAAATGAGCAGTGGAGTAA 302  
403 GluAsnLeuHisArgAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlu 422  
QY ::::: 274  
Db 303 GAGATCTTCATAGGAGTACGAGGAGTACGATCTGATGCTGAGGAGGATGATGAG 362  
423 AspAsnLeuHisSerProLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle 442  
QY ::::: 274  
Db 363 GACATCTCCATAGCCCATTTGCTGCCAGGAGGCAACAGGTGCGGAAGGAAGACATT 422



QY 443 ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlu 462  
 Db 423 GTGCACATGGTCCACCGTGAAGTGTCTTCAGCATGAGAAGGCAAGACCTCTTAGGGGAG 482  
 QY 463 GlyGlyAspGlyValSerSerThrArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 482  
 Db 483 GGTGGAGATGGTGTGAGCAGCACTGATATCGGTGGGGGATGGCAGCTTGTCTGGAATGG 542  
 QY 483 SerGluGlyGlyGlyGluAenGlyArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 502  
 Db 543 TCAGAGAAGAGGCTGAGAATGTTAGAAAGGAGGTGGTTCAAAAGAGTCTACTTGCAC 602  
 QY 503 GlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAsp 522  
 Db 603 CAAGAGGAGTCTCTGGCTCAAGAAAGGAGGTCAATTTGTTCACTCCCGTGGTGGCGAT 662  
 QY 523 ValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPhe 542  
 Db 663 GTTCTTGAGGGTAGTGAGTTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTTTTC 722  
 QY 543 SerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVal 562  
 Db 723 TC-AAGGGTCTGCTGAACA-CGATGTCAGATGCTGCCATGTTTCAACCATCTGAGGTA 780  
 QY 563 AlaAlaLysGlySerArg-TripLysAspLeuPheGluProGlyValArgArgAlaLeuLe 582  
 Db 781 GTGCCAAAGGTTACGTTTGGAAAGATTGTTTGAACCTGGAGTGGAGCGTGCCCTGTT 840  
 QY 582 uVal 583  
 Db 841 AGTC 844

RESULT 6  
 BZ989776/c  
 LOCUS  
 DEFINITION  
 PUGD077D ZM\_0.6\_1.0\_KB Zea mays genomic clone ZMMBTA362B13,  
 genomic survey sequence.  
 BZ989776  
 BZ989776.1 GI:29229913  
 GSS.  
 Zea mays  
 Zea mays  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 795)  
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
 Bennetzen,J.  
 Maize Genomics Consortium  
 Unpublished (2003)  
 Other GSSs: PUGD077B  
 Contact: Cathy Whitelaw  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TF  
 Class: sheared ends.  
 Location/Qualifiers  
 1..795  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBTA362B13"  
 /clone\_lib="ZM\_0.6\_1.0\_KB"  
 /notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
 Cot selected genomic DNA library"  
 Cot

TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source

ORIGIN  
 Alignment Scores:

US-10-051-909-32 (1-800) x BZ989776 (1-795)

QY 270 SerGlyGluLeuSerLeuLeuGlyGlyLeuGluValGlyGlyAspThrSerIleGlu 289  
 Db 738 ACAGGTGAATTTGCTCCCTTCTTCGAAGGTTGGAGGTTGGAGGACACTTCCATTGAA 679  
 QY 290 GluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLys 309  
 Db 678 GAGTACATCATTTGGACCTCCACCGAGGAGCGCATGATCTTGTACTACCGTGATNAAG 619  
 QY 310 GluGlnIleThrLeuTyrGlyProGluGlyGlyGlnSerTyrIleAlaArgProSerLys 329  
 Db 618 GAACAAATCACACTTTATGGGCTGAAGAGGCCAGTCATGGATTGCTCGACCTTCACAG 559  
 QY 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349  
 Db 558 GGACCCAGCATGCTTGGAAAGTGTGCTTCTTTCATCTCGTCATGGAGCATGGTGAAC 499  
 QY 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369  
 Db 498 CAGAGTGTACCCCTTATGATCGATGTCGATGTCACACTTTTGGTAGTGTCCATGAGATATG 439  
 QY 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389  
 Db 438 CCTCAAGCTGGAGGAAGTATGAGGAGCACTTGTTCCTCAAACTTTTGGAAAGTATGTTCACT 379  
 QY 390 ValThrAspGlnHisAlaLysAsnGluGlnTyrAspGluGluAsnLeuHisArgAspAsp 409  
 Db 378 GTCAAGATCAGCATGTCACAAATGAGCAGTGGATGAAGAGAAATCTTCATAGGGATGAC 319  
 QY 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu 429  
 Db 318 GAGGAGTAGCATCTGATGTCAGAGGAGTACTATGAGGAGCAATCTCCATAGCCCATTTG 259  
 QY 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly 449  
 Db 258 CTGTCAGCAGCAGCAACAGTGGGGAAGGAAGACATTTGTCACCATGTTCCACCTGGA 199  
 QY 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyGlyGlyGlyGly 469  
 Db 198 AGTGCTTTGAGCATGAGAAGGCAAGACCTCTTAGGGGAGGTTGGAGATGTTGTGAGCAGC 139  
 QY 470 ThrAspIleGlyGlyGlyTyrGlnLeuAlaTyrLysSerGlyGlyGlyGlyGlyGlyGly 489  
 Db 138 ACTGATATCGTGGGGGATGGCAGCTTGTCTGGAATGGTCAAGAAAGGAGTGAAT 79  
 QY 490 GlyArgGlyGlyGlyGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509  
 Db 78 GGTAGAAAGAGGTTGTTTCAAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTAGCTCA 19  
 QY 510 ArgArgGlySerIleVal 515  
 Db 18 AGAAGGGGCTCAATTTGTT 1

RESULT 7  
 CNS0A3H4  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GSLTSL1502E01 of Silique of strain col-0 of Arabidopsis thaliana  
 (thale cress).  
 BX828912  
 BX828912.1 GI:42461191  
 HTG; GSLT\_cDNA.  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@unicamp.br  
 Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>  
 Plate: 039 row: G column: 07  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1. .833  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCACAD1039G07"  
 /lab\_host="DH10B"  
 /clone\_lib="AD1"  
 /notes="Organ: seedlings inoculated with Gluconacetobacter diazotrophicans; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [seedlings inoculated with Gluconacetobacter diazotrophicans]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

# FEATURES

source

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,77e-107 Length: 833  
 Score: 1168.00 Matches: 244  
 Percent Similarity: 90.65% Conservative: 8  
 Best Local Similarity: 87.77% Mismatches: 26  
 Query Match: 28.66% Indels: 2  
 DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA065405 (1-833)

Qy 487 GlyGluAsnGlyArgIysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal 506  
 Db 2 GGTGAGGATGTTGAAGAAGAGGTGGTTTCAAAAGAGTCTACTACCAAGAGGGAGTT 61  
 Qy 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly 526  
 Db 62 CCTGGCTCAAGAAATGGGCTCAATTTTCTACTTCTCTGGTGGCGATGTTCTCGAGGT 121  
 Qy 527 SerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerIysGlyLeu 546  
 Db 122 GCGAGGTTGTGTACATGCTGCTGTTTAGTAAGCAGTCAGCAGCTTTTCTGAAGGATCTT 181  
 Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566  
 Db 182 ACCGACACCGATGTCGTGTCGCATGTTTACCATCCGAGGTAGTGCACAAAGT 241  
 Qy 567 SerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuValGlyValGly 586  
 Db 242 TCAAGTTGGAAAGATTGTTTGAACCTGGTGTGAGGCGTCCCTGTTAGTCGGTGTGGA 301  
 Qy 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606  
 Db 302 ATTACAGATCTTCAACAGTTTGTGGAAATAAATGGTGTCTGTACTATACCCACAAAT 361  
 Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerAlaSerAla 626  
 Db 362 CTGAGGCACTGGCTGGGAGATCTCTCTTCCAAATTTGGTCTCAGCTCGGATCAGCA 421  
 Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646  
 Db 422 TCCATCTTGATCAGTTCTCTCACTACCTTACTGATGCTTCTAGCATTTGCTTAGCCATG 481  
 Qy 647 LeuLeuMetAspLeuSerGlyArgPheLeuLeuGlyThrIleProIleLeuIle 666

Db 482 AGACTTATGATCTTTCTGGAAGAGGTTTTCCTGCTAGGCACAATTCCAATCTTGATA 541  
 Qy 667 AlaSerLeuValIleLeuValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686  
 Db 542 GCATCTTTAGTTATCTCTGGTCGTGTCCAATGTTTATTGACTTGGGTACAGTGGCCCATGCT 601  
 Qy 687 LeuLeuSerThrValSerValIleValTyrPheCysPheValMetGlyPheGlyPro 706  
 Db 602 GCCTCTCTCCACAGTCAGTGCATCTACTTCTGCTGCTNTGTCTATGGGATTTGGTCCC 661  
 Qy 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726  
 Db 662 ATCCCAACATCTATGTGCAGAGATCTTNTCCACTAGGGTTCGGGTCTCTGCATTTGCC 721  
 Qy 727 IleCysAlaPheThrPheTrpIleGlyAspIleValThrTyrSerLeuProValMet 746  
 Db 722 ATCTGTGCTTGACATNTNNGATCAGACATCATTTGTACCTACAGCCCTCCTGTGATG 781  
 Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeu 764  
 Db 782 CTGAATGCTATTGGGACTAGCAGTGTTTTCG-ATATATGCAGC-GTATGCTTG 833  
 RESULT 9  
 CA146057  
 LOCUS SCVPRT2076D10.g RT2 Saccharum officinarum cDNA clone SCVPRT2076D10  
 DEFINITION 5', mRNA sequence.  
 ACCESSION CA146057.1 GI:35044714  
 VERSION CA146057  
 KEYWORDS EST.  
 SOURCE Saccharum officinarum  
 ORGANISM Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.  
 REFERENCE 1 (bases 1 to 733)  
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
 The libraries that made SUCEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@unicamp.br  
 Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>  
 Plate: 076 row: D column: 10  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1. .733  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCVPRT2076D10"  
 /lab\_host="DH10B"  
 /clone\_lib="RT2"  
 /notes="Organ: Root tips (0.3cm-long) from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Root tips (0.3cm-long) from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript plasmid system kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

## FEATURES

source

## ORIGIN

Alignment Scores:  
 Pred. No.: 8,04e-106 Length: 733  
 Score: 1154.00 Matches: 221  
 Percent Similarity: 93.88% Conservatives: 9  
 Best Local Similarity: 90.20% Mismatches: 14  
 Query Match: 28.32% Indels: 1  
 DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA146057 (1-733)

QY 340 LeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleVal 359  
 Db 1 CTCGCATCTCGTCATGCGAGCATGCTGAACAGAGTGTACCCCTTATGGATCGGATTG 60

QY 360 ThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThr 379  
 Db 61 ACACATTTTGGAGTGTCCAGAGATAGCTCAGCTGAGGAGATGATGAGAGCACA 120

QY 380 LeuPheProAsnPheserMetPheSerValThrAspGlnHisAlaLysAsnGluGln 399  
 Db 121 TTGTTTCCAACTTTGGAGATGTTTCAAGTGTCAAGATCAGCATGCGCAAAATGAGCAG 180

QY 400 TrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGly 419  
 Db 181 TGGGACGAGAGAAATCTTTCAAGGACGATGAGGAGTATGTCATCTGATGTCAGGAGGT 240

QY 420 AspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGly 439  
 Db 241 GATTATGAGGCAATCTCCACAGCCCATTTGCTGCCAGGAGAGCAACCAAGTGTGAAGGA 300

QY 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeu 459  
 Db 301 AAGGACATTTGTCACCATGTCACCGTGAAGTGTCTTGGACATGAGAGGCAAGCCCTC 360

QY 460 LeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAla 479  
 Db 361 TGGGGGAGGCTGGAGAGGAGTGTGAGCAGCACTGATTTGTTGAGAGATGGCAGCTTGA 420

QY 480 TrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgVal 499  
 Db 421 TGGAAATGTCAGAGAGAGAGAGGAGTGAAGTGAAGGAGAGTGTCTCAAAAGAGTC 480

QY 500 TyrLeuHisGlnGluGlyValProGlySerArgGlySerIleValSerLeuProGly 519  
 Db 481 TACTTACACCAAGAGAGGAGTTCCTGGCTCAAGATGGGCTCAATGTTTCACTTCCTGGT 540

QY 520 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSer 539  
 Db 541 GGTGGCAATGTTCTCTGAGGTGGGAGTGTGTATCATGCTGCTGTTTAGTAAGCCGTC 600

QY 540 AlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisPro 559  
 Db 601 GCACATTTTTCGAAGATCTTTACCGAACCCAGCATGTCGTGTCGTCATGTTTCAACCA 660

QY 560 SerGluValAlaAlaLysGlySerArg-TrpLysAspLeuPheGluProGlyValArg 579  
 Db 661 TCCAGGTAGCTGTCGCAAGGTTCAAAGTTGGAAGATTTGTTGAACTGCTGTTGAGGG 720

QY 579 GAlaLeuLeuVal 583  
 Db 721 TGCCCTTGTAGTC 733

RESULT 10  
 CA075335 696 bp mRNA linear EST 23-SEP-2003  
 LOCUS SCUJFAM1069A06.g AM1 Saccharum officinarum cDNA clone SCUJFAM1069A06  
 DEFINITION 5', mRNA sequence.  
 CA075335  
 CA075335.1 GI:34927608  
 EST.  
 Saccharum officinarum  
 Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
 complex.  
 1 (bases 1 to 696)  
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
 The libraries that made SUCEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: paruda@unicamp.br  
 Clone distribution: Clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bcccnet.fcv.unesp.br  
 Plate: 069 Row: A Column: 06  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1..696  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCUJFAM1069A06"  
 /lab\_host="DH10B"  
 /clone\_lib="AM1"  
 /note="Organ: Apical meristem and tissues surrounding of  
 mature plants; Vector: pSport1; Site 1: SalI; Site 2:  
 NotI; An unidirectional cDNA library generated from  
 [apical meristem and tissues surrounding of mature  
 plants]. cDNA was prepared from polyA+ mRNA using  
 SuperScript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 Kb were directionally cloned into the vector. Details  
 of each source of RNA and library construction can be  
 obtained at http://sucest.lad.ic.unicamp.br/public"

Alignment Scores:  
 Pred. No.: 1.21e-104 Length: 696  
 Score: 1142.00 Matches: 214  
 Percent Similarity: 95.22% Conservatives: 5  
 Best Local Similarity: 93.04% Mismatches: 11  
 Query Match: 28.02% Indels: 0  
 DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA075335 (1-696)

QY 305 ThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIle 324  
 Db 1 ACTGATGTCATATAGGAACAAATCACACTTTATGGGCTGAAGAGGCCAGTCATGAT 60

QY 325 AlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHis 344  
 Db 61 GCTGACCTTCCAAGGACCCAGCATGCTTGGAGTGTGCTTCTGTCATCTGTCAT 120

QY 345 GlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySer 364  
 Db 121 GCGAGCATGTGGAACAGAGTGTACCCCTTATGATCCGATGTGACACTTTTGGAGT 180

QY 365 ValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPhe 384  
 Db 181 GTCCACGAGATATGCTCAAGCTGGAGGAGTATGAGGAGCACATGTTTCCAACTTT 240

QY 385 GlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsn 404  
 Db 241 GGAAGTATGTTCAAGTGTCAAGATCAGCATGCCAAAATGAGCAGTGGGACGAGAGAAT 300

QY 405 LeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsn 424  
 Db 301 CTTTCAAGGACCATGATGAGGAGTATGATCTGTATGGTGCAGGAGGTGATTTATGAGGACAAT 360



Db

841 GGAGCTGAAGAGGGGTA-TCTTGGATCGCCCGT-----CTGTTAGG 881

QY

336 SerVal-----LeuSerLeuAlaSerArgHisGlySerMetVal 348

Db

882 CGCGTGGCAAGTGCACATTGGAAGCGCTTGGGTCTCATGTCTCGTATGGAGTATGGTT 941

QY

349 AnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368

Db

942 AGTCAGGGTAA-TCTCTCGTGGACCCACT-GTTACTCTCTTCGGAAGTGCATAGAAGA 999

QY

369 MetProGlnAlaGlySerMetArgSerThrLeuPheProAnPheGlySerMetPhe 388

Db

1000 ----TCCTCGATATGGGAGCATTCGCCCAATTTGTTCTTAATT-GGCCACATGTTT 1055

QY

389 Ser 389

Db

1056 ACG 1058

RESULT 12

BO752737

LOCUS

DEFINITION

WHE4118\_E11\_122ZS Wheat salt-stressed root cDNA library Triticum aestivum cDNA clone WHE4118\_E11\_122, mRNA sequence.

ACCESSION

BO752737

VERSION

BO752737.1

KEYWORDS

EST.

SOURCE

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

REFERENCE

Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K., Dvorak,J., Lazo,G.K., Pham,J., Rausch,C.J., Wilson,C. and Woo,J. The structure and function of the expressed portion of the wheat genomes - Salt-stressed root cDNA library

UNPUBLISHED (2002)

CONTACT: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

Seq primer: SK primer.

FEATURES

source

1..829

Location/Qualifiers

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE4118\_E11\_122"

/tissue\_type="Roots"

/dev\_stage="Full tillering"

/lab\_host="E. coli SOUR"

/clone\_lib="Wheat salt-stressed root cDNA library"

/notes="Vector: Lambda Uni-ZAP XX, excised phagemid pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hydroponic plants grown to full tillering stage were treated with 150 mM NaCl for either 12 hours or 7 days. Root tissues of the plants subjected to both types of treatment were collected separately at University of California, Davis (E. Akhunov and K. Deal in J. Dvorak's Lab). Total RNA was prepared separately from the two samples (12h and 7day treatments), and equal amount of RNA was then pooled. PolyA RNA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak) at the University of California, Davis. Colony plating, plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

2.28e-100

1101.00

89.89%

82.40%

27.02%

5

829

220

25

2

US-10-051-909-32 (1-800) x BO752737 (1-829)

QY

486 GluGlyGluAsnGlyArgLysGlyGlyPheLysArgValThrLeuHisGlnGluGly 505

Db

28 AAAGCGGAGGATGGCAAGAGAGGAGGCTTCAAAAGAACTACTATTGCACCAAGAGGGG 87

QY

506 ValProGlySerArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525

Db

88 GTGGCGGACTCAAGAAAGGGGCTCTGTGTTTTCACCTCTCTGGTGGGGTGTATGCCACCAA 147

QY

526 ----GlySerGluPheValHisAlaAlaValSerGlnSerAlaLeuPheSerIlys 544

Db

148 GGGGGGAGGGGTTTATACACCTCTCTGTTGTAAGCCACTGGCTCTTTTACTTCCNAG 207

QY

545 GlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisProSerGluValAla 563

Db

208 GATCTTATGGAAGAGCGTATGGCGCGGTCAGCCATGATCCATCCATTGGAGCGAGCT 267

QY

564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuVal 583

Db

268 CCCAAGGTTCAATCTGGAAAGATCTGTTGAACCTGGTGTGAGGCGGTGCAATTGTCGTC 327

QY

584 GlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuThrTyrThr 603

Db

328 GGGTTGGAATTCAGATGCTTCAGCAGTTTGTGGAATAAATGGAGTTCTCTACTATACT 387

QY

604 ProGlnIleLeuGluAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623

Db

388 CCTCAATTCGAGCAAGCTGGTGTGGGGTCTCTTCTTCCATCTTGGCTCAGTTCA 447

QY

624 AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGly 643

Db

448 GCATCAGCTCCATCTTGATCAGCTCTCTCACCACCTTACTCATGCTCCCAAGCANTGCT 507

QY

644 PheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIlePro 663

Db

508 GTAGCCCATGAGACTTATGGATATATCTGGAAGAAGGTTCTTCTACTGGGCACAAATCCC 567

QY

664 IleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeu 683

Db

568 ATCTTGATAGCATCCCTAATTTGTTGGTTGTGTGCCAATGTATCAACTGAGTACGGTG 627

QY

684 AlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGly 703

Db

628 CCCACGCTGTGCTCTCCACAGCTAGCGTCAATTGCTACTTCTGCTGTTGTCTATGGGC 687

QY

704 PheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu 723

Db

688 TTGGCCCAATCCCAACATTTATGTGCAGAGATTTTCCCAACAGAGTCCGTGCTGTC 747

QY

724 CysIleAlaIleCysAlaPheThrPheThrIleGlyAspIleIleValThrTyrSerLeu 743

Db

748 TGCATCGTATTTCGCCCTCCATTTCTGGATATGTGACATTTATTGTTTACTACAGACCTG 807

QY

744 ProValMetLeuAsnAlaIle 750

Db

808 CCTGTGATGCTGAATGCTATT 828

RESULT 13

CA170109

LOCUS

DEFINITION

CA170109

871 bp

mRNA

linear

EST 24-SEP-2003

SCQGSB1083B09.g SB1 Saccharum officinarum cDNA clone SCQGSB1083B09

5', mRNA sequence.

CA170109  
CAI70109.1 GI:35093205  
EST.  
Saccharum officinarum  
Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.  
1 (bases 1 to 871)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenhariaia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bccc.net.br  
Plate: 083 row: B column: 09  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1..871  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCQSB1083B09"  
/lab\_host="DH10B"  
/clone\_lib="SBI"  
/note="Organ: Stalk Bark from adult plants; Vector:  
pSport1; Site\_1: SalI; Site\_2: NotI; An unidirectional  
cDNA library generated from [Stalk Bark from adult  
plants]. cDNA was prepared from polyA+ mRNA using  
SuperScript Plasmid System Kit (Invitrogen). The  
double-strand cDNAs were fractionated in a sepharose  
CL-2B 40cm-columns and fragments sizing between 0.8 and  
1.5 Kb were directionally cloned into the vector. Details  
of each source of RNA and library construction can be  
obtained at http://sucstat.lad.ic.unicamp.br/public"

ORIGIN  
Alignment Scores:  
Pred. No.: 3,11e-100 Length: 871  
Score: 1100.00 Matches: 215  
Percent Similarity: 86.11% Conservative: 33  
Best Local Similarity: 74.65% Mismatches: 40  
Query Match: 26.99% Indels: 2  
DB: Gaps: 0

US-10-051-909-32 (1-800) x CA170109 (1-871)  
Qy 420 AsPTyrGluAspAanLeuHisSerProLeuLeuSerArgGlnAlaThrGlyValaGluGly 439  
Db 9 GACTATGAAACAATGTACACAGCCCTCTGCTGTCCGACAGAAATACAAAGTGCCGAGGG 68  
Qy 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeu 459  
Db 69 AAGGCAATTCACATCACGGGACCGTGGAGTGTCTTGGGTTTGAGAGGAAGACGCTG 128  
Qy 460 LeuGlyGluGlyClyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAla 479  
Db 129 TCCGATGAGGGTGGTGAAGGACCAACACATTCGGCATTTGGGGGATGAGGAGCTGCC 188  
Qy 480 TrpLysTrpSerGluLysGluGluAanGlyArgLysGluGlyGlyPheLysArgVal 499  
Db 189 TCGAATGTCTCAGCGAGAGAGGTGAGCGGTAGAGAGGAGGAGGATTTCAAAGGAATC 248  
Qy 500 TyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGly 519

URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 Genome released by MIPS (Munich information center for Protein  
 Sequences).  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source  
 1. .1107  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSUFB382G07"  
 /tissue\_type="flowers and buds"  
 /clone\_lib="Arabidopsis thaliana flowers and buds Col-0"

ORIGIN

Alignment Scores:  
 Pred. No.: 1.19e-98 Length: 1107  
 Score: 1086.00 Matches: 220  
 Percent Similarity: 77.27% Conservative: 52  
 Best Local Similarity: 62.50% Mismatches: 71  
 Query Match: 26.65% Indels: 9  
 DB: 5 Gaps: 4

US-10-051-909-32 (1-800) x BX841312 (1-1107)

Qy 41 SerSerGluGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60  
 Db 24 TCCTGTAATTCGCGGGAATCTACTAAGCAGTTGTTGAGCAATGAGGAGCGACT 83  
 Qy 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80  
 Db 84 CTGTTGCTCTCCGCCCAATCGCAATCTTCTTACAGATGGACAAATGCCACCAT 143  
 Qy 81 AlaAlaValLeuThrIleLysGlyGluPheGlnGlnGlnGluProThr---Val 99  
 Db 144 GCTGGAGCTATGTTTATCAACAAAGCTTGAATCTA-----CCAACCTCTGTT 194  
 Qy 100 GluGlyLeuIleValSerMetSerIleGlyAlaThrIleValThrPheSerGly 119  
 Db 195 CAGGCTTGTGCTGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 254  
 Qy 120 ProLeuSerAspSerIleGlyArgArgProMetLeuLeuLeuLeuLeuLeuLeuLeu 139  
 Db 255 CCGATATCTGATGCTGCGGAGCGCCCAATGCTCATTTTATCATCATGTTATGTTTC 314  
 Qy 140 PheSerGlyLeuLeuMetLeuTrpSerProAsnValTrpValLeuLeuAlaArgPhe 159  
 Db 315 GTCTGCGTTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 374  
 Qy 160 ValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTrpIleSerGluIle 179  
 Db 375 CTAAATGGTTTGGTCCGCGCTCGCGTTACACTTGTCCCTGTTTACATTTCTGAAACC 434  
 Qy 180 AlaProSerGluIleArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 199  
 Db 435 GCTCTCCGAGATCAGAGACAGTAAATACTCTCCCTCAGTTCTTGGCTCTGTTGGA 494  
 Qy 200 MetPheLeuSerTrpCysMetValPheGlyMetSerLeuSerProSerProAspTrpArg 219  
 Db 495 ATGTTTGTGCTACTGATGTTGTTTCACTATGCTCCCTGAGTACCTCCCTAGCTGGAGA 554  
 Qy 220 IleMetLeuGlyValLeuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTrp 239  
 Db 555 GCCATGCTCGGTGCTCTCGATCCCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 614  
 Qy 240 LeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysVal 259  
 Db 615 TTGCCCGAGTCTCTCGTTGGTGGTTAGTAAAGGAAGATGGACGAGGCTAAGCGAGTT 674  
 Qy 260 LeuGlnLysLeuArgGlyLysAspValSerGlyLeuSerLeuLeuLeuGly 279

Db 675 CTTCAACAGTTATGTGGCAGAGAAGATGTTACCGATGAGATGGCTTTACTAGTTGAAGGA 734  
 Qy 280 LeuGluValGlyGlyAspThrSerIleGluGlyTrpIleLeuGlyProAlaThrGluAla 299  
 Db 735 CTAGATATAGGAGGAGAGAAAAACAATGAGATCTCTTAGTAACCTTTGGAGGATCATGAA 794  
 Qy 300 AlaAspAspLeuValThrAspGlyAspLysGlu---GlnIleThrLeuTrpGlyProGlu 318  
 Db 795 GGTGATGATACACTTGAACCCGTTGATGAGGATGCAAAATGCGGCTTTATGGAACCCAC 854  
 Qy 319 GluGlyGlnSerTrpIleAlaArgPro-SerLysGlyProIleMetLeuGlySerVal 338  
 Db 855 GAGATCAATCGTACCTTGTAGACCTGTCACAGAACAAATA-----GCTCACTT 905  
 Qy 338 uSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspPro 358  
 Db 906 GGGGCTACGCTCTCGCCACGGAAGCTTAGCAACCCAAAGCATGATCTCTTAAAGATCCGCT 965  
 Qy 378 rThrLeuPheProAsnPheGlySerMetPheSer 389  
 Db 1026 TGGGATTTTCCAAACATTTTCGGGAGCATGTTTCAGT 1059  
 RESULT 15  
 CF445664 848 bp mRNA linear EST 04-SEP-2003  
 EST682009 normalized cDNA library of onion Allium cepa cDNA clone  
 ACAIL96, mRNA sequence.  
 CF445664  
 CF445664.1 GI:34468366  
 EST.  
 KEYWORDS  
 SOURCE Allium cepa (onion)  
 ORGANISM Allium cepa  
 REFERENCE 1 (bases 1 to 848)  
 AUTHORS Havey, M.J., Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.  
 TITLE Expressed Sequence Tags from a normalized library of mixed onion tissues (Allium cepa)  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Havey MJ  
 Department of Horticulture  
 USDA-ARS and University of Wisconsin  
 1575 Linden Drive, Madison, WI 53706, USA  
 Tel: 608-262-1830  
 Fax: 608-262-4743  
 Email: mjhavey@facstaff.wisc.edu  
 TIGR sequence name ACAIL96TR. For more information:  
 http://haveylab.hort.wisc.edu  
 Seq primer: CAG GAA ACA GCT ATG ACC.  
 Location/Qualifiers  
 source 1. .848  
 /organism="Allium cepa"  
 /mol\_type="mRNA"  
 /cultivar="Red Creole (bulbs), unknown (callus), Ebano & Texas Legend (roots)"  
 /db\_xref="taxon:4679"  
 /clone="ACAIL96"  
 /tissue\_type="Callus, roots, and young bulbs"  
 /clone\_lib="normalized cDNA library of onion"  
 /notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1: EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA from callus, roots, and young bulbs were combined to synthesize the library. Normalization to enrich for low-copy transcripts was performed by proprietary techniques of Invitrogen."

ORIGIN

Alignment Scores:





This Page Blank (uspro)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:34:09 ; Search time 5504.66 Seconds

(without alignments)  
4867.828 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779

Sequence: 1 PSSSSFRPAGKKKKKNG.....TIFVLSIORQLWLPECLS 553

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0			0.5
Ygapop 6.0			7.0
Delop 6.0			7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp  
-O=/cpn2\_1/USPTO.spool\_p/US10051909/runat\_13042005\_074037\_14048/app.query.fasta\_1.1678  
-DB=GenEmbl -QPWT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10051909 @CGN 1.1 9091 @runat\_13042005\_074037\_14048 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sta.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2133	76.8	1870	8 AK100792	AK100792 Oryza sat
2	2017.5	72.6	2266	8 AK065497	AK065497 Oryza sat
3	1821.5	65.5	1948	8 AK068594	AK068594 Oryza sat
4	1757.5	63.2	2020	8 BVU43629	U43629 Beta vulgar

5	1746.5	62.8	1557	8	AY093274	Arabidops
6	1746.5	62.8	1866	8	AY059848	Arabidops
7	1738.5	62.6	1695	8	AF412060	Arabidops
8	1734.5	62.4	1464	8	AY124845	Arabidops
9	1726.5	62.1	1914	8	AF462803	Arabidops
10	1710	61.5	1377	6	AX653119	Sequence
11	1693	60.9	1825	8	BT014234	Lycopersi
12	1577	56.7	1428	6	AX653120	Sequence
13	1453.5	52.3	204649	8	AC120986	Oryza sat
c 14	1137.5	40.9	120977	8	AC025808	Genomic s
c 15	1117.5	40.2	68041	8	AC025814	Arabidops
c 16	992.5	35.7	99587	8	AC007858	Oryza sat
c 17	982	35.3	156643	8	AC120988	Oryza sat
18	965.5	34.7	1699	8	LES278765	Lycopersi
19	954.5	34.3	2008	8	AK106658	Oryza sat
20	951.5	34.2	1741	8	AY084684	Arabidops
21	944.5	34.0	1778	8	AY064144	Arabidops
22	943.5	34.0	1435	8	BT000053	Arabidops
23	943.5	34.0	1686	8	AY120715	Arabidops
24	939	33.8	1449	8	BT000608	Arabidops
25	935.5	33.7	1747	8	AY087180	Arabidops
26	932	33.5	1806	8	BT008661	Arabidops
27	932	33.5	1847	8	ATH249967	Arabidops
28	920.5	33.1	1473	6	AX507636	Sequence
29	920.5	33.1	1678	8	D89051	Arabidops
30	899	32.3	1359	6	AX506932	Sequence
31	898.5	32.3	1420	8	AY091216	Arabidops
32	898.5	32.3	1660	8	AY063856	Arabidops
33	891.5	32.1	1755	8	AK064996	Oryza sat
34	878.5	31.6	1729	8	AF367260	Arabidops
35	877.5	31.6	1407	6	AX506596	Sequence
36	877.5	31.6	1407	6	AX651294	Sequence
37	877.5	31.6	1413	8	AY133547	Arabidops
38	870.5	31.3	1606	8	AY026255	Arabidops
39	866	31.2	1578	8	ATH249968	Arabidops
40	858.5	30.9	1757	8	AY048207	Arabidops
41	858.5	30.9	1757	8	AY123991	Arabidops
42	853.5	30.7	1787	8	AY026254	Arabidops
43	845	30.4	1668	8	AY058056	Arabidops
44	843.5	30.4	1377	8	BT002699	Arabidops
45	837.5	30.1	1209	6	AX653566	Sequence

#### ALIGNMENTS

AK100792	AK100792	1870 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	AK100792	Oryza sativa (japonica cultivar-group)	cdna clone:J023121B20, full		
DEFINITION	AK100792	Oryza sativa (japonica cultivar-group)	insert sequence.		
ACCESSION	AK100792	AK100792.1	GI:32986001		
VERSION	AK100792	FLI CDNA: CAP trapper.			
KEYWORDS	AK100792	Oryza sativa (japonica cultivar-group)			
SOURCE	AK100792	Oryza sativa (japonica cultivar-group)			
ORGANISM	AK100792	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			



```

Qy 314 SerSerSerLysArgThrThrIleSerPheGlnGluLeuAenGlnLysLysTyrArgThr 333
Db 999 TCAGCAAAACAAAGGACACAGCATCCGTTTTCAGAGATTAACACCAAGAAATACCGCAC 1058
Qy 334 ProLeuLeuGlyIleGlyLeuValLeuGlnAenLeuSerGlyIleAenGlyVal 353
Db 1059 CCCTAATATAGGAATGGCTACTTGTACTGCACAGCTAAGTGAATCAATGGGAATA 1118
Qy 354 LeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAenSerAspLeuAlaThr 373
Db 1119 TTGTTTTATGCAGTAGCATCTTCAAAGCAGCAGGTCTCAAAACAGTGAGCTTGGCTACA 1178
Qy 374 CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArg 393
Db 1179 TGTGCATCTGGTGTATCCAGGTTCTTGTACAGGAGTTCAACCTGGTTATTAGACAGA 1238
Qy 394 AlaGlyArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAla 413
Db 1239 GCTGGCCGAGGATCCCTTATCATCTCTCTGCTGGGATGACTCTAAGCCTCTTGCA 1298
Qy 414 ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAenSerTyrTyrIle 433
Db 1299 GTTGCTGTGTATTTTCTCAAGGATAGCATTTCAAGATTTCTCACATGTACTACACC 1358
Qy 434 LeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db 1359 TTAAGTATGATCTCTCTGGTGTCTTGTGTGCTTTTGTAAATCGCTTCTCTCGGTATG 1418
Qy 454 GlyValIleProThrLeuMetMetSerGluLeuLeuProValSerIleLysSerLeuGly 473
Db 1419 GGTGCCATTCATGATCAATATGTACAGATCTCCCGGTGTAGTATCAAGAGTCTCGCA 1478
Qy 474 GlySerIleAlaThrLeuAlaAenThrLeuThrSerPheAlaIleThrMetThrThrAsn 493
Db 1479 GGAAGCTTTGGAGCGCTCGCCAACTGGCTTACATCTTTTGGATAACAATGACAGCAAC 1538
Qy 494 LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
Db 1539 TTGATGCTTACGTGGAGTGTGGAGGACCTTTGTGTCTTACATGTCGTGAGTGCCTTC 1598
Qy 514 ThrIleValPheValValLeuThrTrpValProGluThrLysGly***AenSerArgGlyAsp 533
Db 1599 ACCCTCGTGTGTGTCATCTTTGGGTGCCAGACAAAGGAAG-AACTTCGAAGAGAT 1657
Qy 534 ThrIlePheValSerLeuSerIle 541
Db 1658 ACAAATGCTCTTCGCTGAGCCCTT 1681

RESULT 2
AK065497
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J013023P09, full
DEFINITION insert sequence.
ACCESSION AK065497
VERSION AK065497.1 GI:32975515
KEYWORDS FULI_CDNA; CAP trapoper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
1. The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team,
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Tada,Y., Sugano,S., Fujimura,T., Masuda,H., Kobayashi,M., Mizuno,K., Niikura,J.,
Kurokawa,R., Sugiyama,A., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S.,
Niikura,J.,
Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashizume,W., Hayatsu,N., Imotani,K., Iishi,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Oeato,N., Oca,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
JOURNAL MEDLINE
22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 2266)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hirooka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numaaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Oeato,N., Oca,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S., and
Yoshimura,A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:shikuchi@nias.affrc.go.jp)
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cdna/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hirooka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numata,K., Sakazume,N., Sano,H.,
Oca,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1. .2266
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
FEATURES
source

```

source





QY	158	GlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIle	177
DB	473	.....	532
QY	178	IleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeu	197
DB	533	ATGGGTGGCTCGCAATATCTTTGGCAAGAGACTCATCTTTTGTATGGCGCGATG	592
QY	198	LeuGluGlyPheGlyValGlyIleIleSerTyrThrValproValTyrIleAaIle	217
DB	593	CTGGAGAGATTGGGGTCGGGTGGGTATCTATGTGTCGACAGTTTATATAGCAAAATT	652
QY	218	SerProGlnAsnMetArgGlyAlaLeuGlyGlySerValAsnGlnLeuSerValThrPheGly	237
DB	653	GETCCACAAACAATGAGAGGAGCTTTGGTTCAGTCAATCAGCTTTCGTACTATTGGC	712
QY	238	IlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGly	257
DB	713	ATATTGCTTCCTACTTCCTTAGGCATCTTTGTTCCCTCGGAGAATTCTCTCCGTTCTCGGT	772
QY	258	AlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyr	277
DB	773	ATTTTACCATGTTCAATTCTCATACCTGGATCTGTTCTTTATCCCTGAATCCACCAAGGTGG	832
QY	278	LeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe	297
DB	833	CTGGCAAAAATGGGAAGATGGAGATTGTTGAATCCTCGCTCAAGTATTGCGAGGATTT	892
QY	298	GluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLys	317
DB	893	GAACACAGATCCGCTGTAGAAATTAATGAATTAAGAGAACAGATTCAATCATCATCGGAGG	952
QY	318	ArgThrThrIleSerPheGlnGluLeuAsnGlnIleLysTyrArgThrProLeuLeuLeu	337
DB	953	AGACGACAAATACGATTTTCGAGATATACAGCAGAGAGATATAGTGATCCTCTAATGATA	1012
QY	338	GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAla	357
DB	1013	GGAAATTGGTCTCCTTGTACTCGACAGTTAAGTGGTGAATGGACATCTATTATTATGCT	1072
QY	358	SerSerIlePheLysAlaAlaGlyValThrAsnSerAspIleAlaThrCysSerLeuGly	377
DB	1073	GCTAGTATCTTCAAAAGCCGCGCTTTACGAAATAGTAATCTAGCAACATTTGGTCTGGGG	1132
QY	378	AlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArg	397
DB	1133	GTTGTTCAGTGGTGTCTACTCGAGTGCACAACTGGTTGACTGACAAAGCTGGTCGAGA	1192
QY	398	IleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVal	417
DB	1193	CTCCTCTCATTTATCTTACTACAGGATGACCATTACTCTTGTTGTTGTTCTGTGTCA	1252
QY	418	PhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIle	437
DB	1253	TTTTTTGTGAGGACACATACATATATGGTTCATTTATCTCTGTATAGATATGCTT	1312
QY	438	SerLeuValGlyIleValSerPheValIleThrPheSerPheClyMetClyAlaIlePro	457
DB	1313	TCGCTGGTCGGGCTGTGGCATTTGTGATTTTCATTTTCTCTTGTTGGGACCATACA	1372
QY	458	TrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAla	477
DB	1373	TGGATCATTAATGCTGTGAGATTTCTTCGTGAATAATCAAGAGCTTCTGCTGGAGAGCTTGCA	1432
QY	478	ThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThr	497
DB	1433	ACCCTTGCAAACTGGTTGACCGCTTGGCTCATTTACGATGACACCAAGCTTGATGTAAGC	1492
QY	498	TrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPhe	517
DB	1493	TGAGCAATGGAGGAACATTTGCTATCTATCTCTCTGTGTGTCGGGACCTCGTCTTTC	1552

Qy	518	ValValleuTrpValProGluThrIysGly****AsnSerArgGlyAspThrIlePheVal	537
Dd	1553	GTATCGCTGTGGTGC CGGAGACTAAAGAG- AACACTTGAGGAATTCGATTCTTT	1611
Qy	538	SerLeuSerIleGlnArgGlnLeu	545
Dd	1612	CCGGTGAAGCTCAGGCCAACGTTG	1635
RESULT	4		
LOCUS	BVU43629	2020 bp mRNA linear	PLN 11-JUN-2001
DEFINITION	Beta vulgaris integral membrane protein mRNA, complete cds.		
ACCESSION	U43629		
VERSION	U43629.1	GI:1209755	
KEYWORDS	Beta vulgaris		
SOURCE	Beta vulgaris		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Beta.		
REFERENCE	1. (bases 1 to 2020)		
AUTHORS	Chiou,T.-J. and Bush,D.R.		
TITLE	Molecular cloning, immunochemical localization to the vacuole, and expression in transgenic yeast and tobacco of a putative sugar transporter from sugar beet		
JOURNAL	Plant Physiol.	110 110 (2), 511-520 (1996)	
MEDLINE	96351183		
REFERENCE	2. (bases 1 to 2020)		
AUTHORS	Chiou,T.-J. and Bush,D.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-Dec-1995) Jen-T. Chiou, Plant Biology, University of Illinois, 1201 W.C. Gregory, 190 ERLM, Urbana, IL 61801, USA		
FEATURES	Location/Qualifiers		
source	1..2020		
	/organism="Beta vulgaris"		
	/mol_type="mRNA"		
	/db_xref="taxon:161934"		
	/cell_type="tonoplast"		
	/note="vacuole membrane"		
CDS	190..1662		
	/function="putative sugar transporter"		
	/note="member of major facilitator superfamily"		
	/codon_start=1		
	/product="integral membrane protein"		
	/protein_id="AAB53155.1"		
	/db_xref="GI:1209756"		
	/translation="MSDSEALGFGGGDLRKPFLHTGSWVRMGSSROSLIMGSSQIVTR ESSISLVACLVI VALPIQDFGTGYASTQTSAITNELGLSVAEYSWFGLSNVGAMV GAIASGAI SEGRKSGLMIAIPNLIGWLASFADKSSFLYVMGLEFGGVGISVT VPVVISIA FQNIKRGALGNOLSVTIGFYMLSMLGLEVPWRILAVLIGLPCTIIIPAE LFPTPEPRNLAKGMWEFEFSLOVLGFDPDISLEVNEIKRSVAASSKRTTIIRPAE LQQRYTWLPMIGNLLLIQLQSLNGVLFPYSTTKEAGVTSNAATFGLGAVQVILA LRVVTLWDVSGRWRLLI VSSSSGSSLLLVAMSFFLKMYSDSTSWVSFPLSVVG WAWMTFTSGIGAIPWIMSELLPINIKGLAGSATLANFWFAWIWTMTANIMLSNW SGGTFSIYWVCAPTFAVVVIVWPETKGRILEEIOWSPRR"		
ORIGIN			
Alignment Scores:			
Pred. No.:	4,56e-122	Length:	2020
Score:	1757.50	Matches:	343
Percent Similarity:	81.87%	Conservative:	77
Best Local Similarity:	66.86%	Mismatches:	81
Query Match:	63.24%	Indels:	13
DB:	8	Gaps:	3
US-10-051-909-36 (1-553) x BVU43629 (1-2020)			
Qy	39	MetGlyGlyGlySerAsnArgGlyGlyAlaGlyValaglyGluGluserGlyserAsphis	58
Dd	190	ATGAGTTCAGATTCAGACGAGGTTTAGGAGGAGGAGGA-----	228
Qy	59	AspGlyValLeuArgArgProLeuLeuAsnThrGlySerTripyrArgMetSerSerArg	78



Db 229 ---GGAGATCTGCGGAACCACTTTTGCATACAGGAAGTTGGTATAGAAATGGGTTCAGA 285  
 Qy 79 GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98  
 Db 286 CAATCAAGTTTG-----ATGGGTCTCTCAAGTTATTCAGAAAGCTCCATCTCT 336  
 Qy 99 AlaPheLeuCythrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
 Db 337 GTCTTGTCTGTCTTATTATTGTCTTGGTCTTATTCAGTTTGGTTTACTGTGGT 396  
 Qy 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138  
 Db 397 TATTCACTCCAATCACTCAGCAATTCACAAGTGAACCTTGTCTTCTGTTCAGAGTAC 456  
 Qy 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158  
 Db 457 TCATGGTTTGGTCTCTATCCAAATGTGGTGCCATGTTGGGGCAATTCGAAGTGGTCAA 516  
 Qy 159 MetAlaGluThrIleGlyArgGlyGlySerLeuMetIleAlaIleProAsnIleIle 178  
 Db 517 ATATCTGAGTATATTGGAAAGAAAGGGTCTTTGATGATTGCTATCCCCAACATAATT 576  
 Qy 179 GlyTriPheAlaIleSerPheAlaIleAspAlaSerPheLeuThrMetGlyArgLeuLeu 198  
 Db 577 GGATGGCTGCTATATCATTTGGTCCAAAGATCTTCATTTCTGTATATGGAGAGATGTTG 636  
 Qy 199 GluGlyPheGlyValGlyIleIleSerThrValProValThrIleAlaGluIleSer 218  
 Db 637 GAAGGTTTGGCGTGGGAATAATATCTTACACGGTCTCTGTGTACATTTCTGAAATAGCA 696  
 Qy 219 ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238  
 Db 697 CCACAGAATTTGAGGGGAGCTTTGGGTTCAGTAACACAGATATCCGTACAAATCGGTATA 756  
 Qy 239 PheLeuAlaThrLeuLeuGlyMetPheIleProTriPheArgLeuLeuAlaValIleGlyAla 258  
 Db 757 ATCTATCGTATATGTCGGGGCTTTTTCGGTGGAGGATCTTGCAAGTTTATAGGAATT 816  
 Qy 259 LeuProCythrMetLeuIleProGlyLeuPheIleProGluSerProArgTriPheLeu 278  
 Db 817 TTACCATGCATATATTGATACCTGGCTTATTTCTTATACCGGAATCTCTAGGTGGTTG 876  
 Qy 279 AlaIleMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298  
 Db 877 GCGAAAATGGGTATGATGGAAGAATTTGAGGTTCCTTGCAGAGTTCTCCGGGGCTTTGAC 936  
 Qy 299 ThrAspIleThrThrGluValAsnAspIleIleArgAlaValAlaSerSerSerLysArg 318  
 Db 937 ACTGATATATCTCTTGAAGTTAATGAATCAAGAGATCTGTGCTTCNAGTAGCAAGA 996  
 Qy 319 ThrThrIleSerPheGlnGluLeuAsnGlnLysLysThrArgThrProLeuLeuLeuGly 338  
 Db 997 ACCACCATAGATTGTCAGAACTACGACAAAGAAGATATTGGTTACCTTGGATAGG 1056  
 Qy 339 IleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheThrAlaSer 358  
 Db 1057 AATGGTTTACTTATTTTGCACCAACTCAGTGGGAATCAATGGTGTCTGTCTTATTCAGT 1116  
 Qy 359 SerIlePheLysAlaIleGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyVala 378  
 Db 1117 ACCATCTTTAAGAGGCTGGGTACATCAACGACCTGCAACATTTGGCCCTTGGTGTCT 1176  
 Qy 379 IleGlnValLeuAlaThrGlyValThrThrPheLeuLeuAspArgAlaGlyArgArgIle 398  
 Db 1177 GTTCAGGTATTCGCAACTGTTGTGACTACCTGTTGGTGGATAAATCAGGACGCAGACTT 1236  
 Qy 399 LeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhe 418  
 Db 1237 CTCTTAATGCTCTCTCTCTGGGATGACATGATGCTTCTAGTGGTGCGAATGTCCTTTC 1296  
 Qy 419 PheLeuLysAspAsnIleSerGlnAspSerAsnSerThrTyrrIleLeuThrMetIleSer 438  
 Db 1297 TTTTAAAGGAATGGGTATCAGACGAATCTACATGGTACAGTGTATTCTCCATCCTTTCA 1356

Qy 439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTriP 458  
 Db 1357 GTTGTGGAGTTGTGGCTATGTGGTTACTTCTCACTAGGAATGGAGCAATTCGGTGG 1416  
 Qy 459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr 478  
 Db 1417 ATTATATGTTCAGAGATTCTTCCAAATTACATCAAGGGCTGGCTGGAGTATTGTCTACA 1476  
 Qy 479 LeuAlaAsnTriPheLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTriP 498  
 Db 1477 CTTCGCAACTGGTTGTAGCATGGATTGTGACAAATGACTGCACAAATAATGTTGTCATGG 1536  
 Qy 499 SerValGlyGlyThrPheLeuSerTyrrMetValValSerAlaPheThrIleValPheVal 518  
 Db 1537 AATAGTGGAGGTACTTCTCTATTATTCATGGTTGTTGTGCTTCTACTGTGGCCTTTGTG 1596  
 Qy 519 ValLeuTriPheValProGluThrLysGly\*\*\*AsnSerArgGlyAspThrIlePheValSer 538  
 Db 1597 GTCATTTGGTCCCGAGACAAAGGGAG-AACATTGGAGGAGATTCAGTG---GTCCTT 1652  
 Qy 539 LeuSerIleGlnArgGlnLeuGlnTriPheLeuProGluCys 551  
 Db 1653 CAGAAGATAAAGAAATTTCTGATAACTGTGCCATGGTGT 1691  
 RESULT 5  
 AY093274  
 LOCUS  
 DEFINITION Arabidopsis thaliana similar to integral membrane protein  
 (Atg19450) mRNA, complete cds.  
 ACCESSION AY093274.1 GI:20259851  
 VERSION AY093274.1  
 KEYWORDS FLI CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 REFERENCE 1 (bases 1 to 1557)  
 AUTHORS Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,  
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,  
 Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,  
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Naruseaka,M., Pham,P.K.,  
 Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,  
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-MAR-2002) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN  
 Arabidopsis Full-length cDNA'): Seki M., Narusaka M., Ishida,J.,  
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
 Hayashizaki,Y. and Shinozaki,K.  
 The Salk, Stanford, PGC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAPL cDNAs: Southwick,A.,  
 Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,  
 Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung M.K.,  
 Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,  
 Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed  
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.  
 (SSP/Stanford) contributed equally to this work as PIs.  
 FEATURES  
 Location/Qualifiers  
 1..1557  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"

/clone="U16352"  
 /ecotype="Columbia"  
 /notes="This clone is in pENTR/SD-dTropo. This is a cloned  
 PCR product using RIKEN clone RAFL07-13-K02 (AY059848) as  
 a template"  
 1. .1557  
 /gene="At1g19450"  
 /notes="synonym: F18014.22"  
 1. .1467  
 /gene="At1g19450"  
 /codon start=1  
 /product="similar to integral membrane protein"  
 /protein\_id="AA13273.1"  
 /db\_xref="GI:20259852"  
 /translation="MSFRDNTTEGRNDRPFLHSGWYRMGRSQSMLESQVIRD  
 SSIIVLACVLIIVAGPIORGTGYSSTQAATIKDLGLTVSEIYSGLSNVGVAVG  
 AIAAGQIAEYVGRKSLMAIPNIIWMLISFPAKDTISFLMGLLGGFGVGLISYIV  
 PUYIARATQWRGLSGVNSQUSITIGIMLAYLLGLFVPMRLAVLGVLPCTLLIPLG  
 FEIPESPRILAKMGLTDDFETISLQVIRGFEDITIVNEIKRSVASSKRSARVFDL  
 KRRIYFFLMVIGIIGLLAQQLGNGVLFITIFESAGVTSNVATFGVGVQVAVT  
 GIATLVADKAGRRLLMISIGMTISLVIVAVFYLFKFEVPSDSNNYILNLSVYGV  
 VANVISCSDMGPIWLMSEILPNIKLAGSITLNNFVSLVMTMTANMLAWS  
 GGTFTYALVCGFTVVFVSLWVPETKLTLEIQALPR"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,12e-121 Length: 1557  
 Score: 1746.50 Matches: 338  
 Percent Similarity: 83.40% Mismatches: 69  
 Best Local Similarity: 69.26% Indels: 8  
 Query Match: 62.85% Gaps: 2  
 DB: 8

US-10-051-909-36 (1-553) x AY093274 (1-1557)

QY 52 GluGluSerGlySerAspHisAlaValLeuArgArgProLeuLeuAsnThrGlySer 71  
 DB 25 GAAGAGGGAAGAATGAT-----CTTGGGACCGTCTTACACAGGAAGT 72  
 QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91  
 DB 73 TGGTATCGGATCGGTTCCAGACCAATCTAGTAG-----TTGGAATCGTCTCAAGTT 123  
 QY 92 LeuArgGluSerHisValSerAlaPheLeuLeuCysThrLeuLeuValAlaLeuGlyProile 111  
 DB 124 ATTCGAGACAGCTCAATTTCTGCTTAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183  
 QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131  
 DB 184 CAATTCGGATTCACCTTGTGGTTATCTCTCAACCACTCAAGCTGCCATTAAGGATCTT 243  
 QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal 151  
 DB 244 GGTTTAACTGTATCAGAGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 303  
 QY 152 GlyAlaIleAlaSerGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171  
 DB 304 GGAGCAATTCGAGT 363  
 QY 172 AlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPhe 191  
 DB 364 GCTGCAATTCCTCAATATTATTGATGCTTTCGATATCATTTGCAAAAGATCTCTTTT 423  
 QY 192 LeuTyrMetGlyArgLeuLeuGlyPheGlyValGlyIleIleSerTyrThrValPro 211  
 DB 424 CTTTACATGGGAAGATTTGTAGAAGGCTTCGGAGTTGGAGTATATCTCTATACGGTCCC 483  
 QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231  
 DB 484 GTATATATAGCTGAGTACGCTCCACAGCATCAGAGGAGCCCTTAGGTTCACTTAACAG 543  
 QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251

Db 544 CTTTCTGTAAACAATTGGGATAATGCTGGCGTATTATTACTCGGTCTCTTTGTTCATGAGA 603  
 QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271  
 Db 604 ATTCITGCAATTTGGGAGTATTGCCATGTACATTTGTATACCGGTCTTTTTCATT 663  
 QY 272 ProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291  
 Db 664 CTGAATCTCTCGGTGTGTGCAAGATGGTGTGACAGATGATTTTCGAAATTCATTG 723  
 QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311  
 Db 724 CAAGTCTCTCGGATTTGAGACTGATATTACCGTTGAGGTATGAAATCAAGAGATCT 783  
 QY 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331  
 Db 784 GTGGCATCATCTAGCAACAGTCTTCGAGTTCGGTTCGTAGACCTCAAGCGCAGAGATAC 843  
 QY 332 ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351  
 Db 844 TATTTCCCACTGATGGTGTGTATAGGCTGCTTGCACCTTCAACAACTTGGAGGAATCAAT 903  
 QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371  
 Db 904 GGTGTCTTGTCTTATTCGAGTACAAATTTTGAATCTGCAGGGGTGACATCAAGTAATGTG 963  
 QY 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrLeuLeu 391  
 Db 964 CGGACATTTGGAGTTCGGGTGTTCAGGTAGTGGGACTGGGATCGGACATGGTGTGGTG 1023  
 QY 392 AspArgAlaGlyArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeu 411  
 Db 1024 GATAGACAGTGTGCGCTTCTACTCATGATCTCTTCTATCGGAATGACGTAAGCCCTC 1083  
 QY 412 LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431  
 Db 1084 GTGATGTAGCAGTGCATTTTACCTTAAGGAATTTGTATCACTCACTGATTTCCAACTGATC 1143  
 QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451  
 Db 1144 AACATCTTAAGCATGGTTCCTGAGTGTGGAGTGTGGTGTGTGTGTGTGTGTGTGTCTCTA 1203  
 QY 452 GlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleLysSer 471  
 Db 1204 GGAATGGGACCAATTCATCGCTGATTTATGTCTGAGATTTCTGCCAGTGAATATAAGGCT 1263  
 QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThrMetThr 491  
 Db 1264 TTAGCCGGAAGTATAGCAACTTTTGTAAACTTGTTCGTGTCTATGTTAGTGAATGACT 1323  
 QY 492 ThrAsnLeuMetLeuThrThrSerValGlyGlyThrPheLeuSerTyrMetValSer 511  
 Db 1324 GCAATATGCTGTAGCATGGAGCAGTGGGGAACCTTTCACTCTACGCTTGTGTGTGTGT 1383  
 QY 512 AlaPheThrIleValPheValValLeuThrValProGluThrLysGly\*\*\*AsnSerArg 531  
 Db 1384 GGATTCACGT 1442  
 QY 532 GlyAspThrIlePheValSerLeu 539  
 Db 1443 AGAGATCCCAAGCTTTGTTCAGATG 1466

## RESULT 6

AY059848

LOCUS

DEFINITION

(At1g19450; F18014.22) mRNA, complete cds.

ACCESSION

AY059848

VERSION

AY059848.1

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

AY059848 1866 bp mRNA linear PLN 04-NOV-2001  
 Arabidopsis thaliana similar to integral membrane protein  
 (At1g19450; F18014.22) mRNA, complete cds.

AY059848

GI:16648956

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1866)

## REFERENCE AUTHORS

Nguyen, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Banh, J., Carninci, P., Chen, H., Chung, M. K., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S. X., Narusaka, M., Pham, P. K., Shinohara, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinzaki, K., Eckert, J., Theologis, A. and Davis, R. W.

Submitted (22-OCT-2001) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

USA  
e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Havareshizaki, Y., and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J. J., Theologis, A. and Davis, R. W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

**FEATURES**  
**SOURCE**

```
1. 1866
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="I"
/clone="RAFLO7-13-K02"
/note="This clone is in pBluescr
```

gene 1. .1866 /mole#1

gene="At1g19450; F18O14.22"

```
249. 1715
/gene="Atg19450; F18014.22"
/codon_start=1
/product="similar to integral membrane protein"
/protein_id="AAL24330.1"
/db_xref="GI:16648955"
```

[illegible]

## ORIGIN

Alignment Scores:		
Pred. NO.:	2.73e-121	Length:
Score:	1746.50	Matches:
Percent Similarity:	83.40%	Conservative:
Best Local Similarity:	69.36%	Mismatches:
Query Match:	62.85%	Indels:
DB:	8	Gaps:
		2
		1866
		338

US-10-051-909-36 (1-553) x AY059848 (1-1866)

QY	52	GluGluSerGlySerAspHisnspGlyValLeuArgProLeuLeuAsnThrGlySer	71
		:	
DB	273	GAAGAGGAAGAATCAT-----CTTCGGCACCCTTTACACACAGGAAGT	32
		:	
QY	72	TyrArgMetSerSerArgGlnSerPheAlaProGlyThrSerSerMetAlaVal	91

Db 1392 AACATCTTAAGCATGGTTTCCGTAGTGGAGTTGGGTATGTTATTTCTTGCTCTCTA 1451  
 QY 452 GlyMetGlyAlaIleProTriPLeuMetMetSerGluLeuLeuProValSerIleLeuSer 471  
 Db 1452 GGAATGGGACCAATTCATCGGTGATGATGTCGTAGATTCGCCAGTGAATATAAAGGGT 1511  
 QY 472 LeuGlyGlySerIleAlaIleThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491  
 Db 1512 TTAGCCGGAAGTATAGCAACTTTGTAAACTGGTTCGTGTCATGGTTAGTACAACTGACT 1571  
 QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTrpMetValValSer 511  
 Db 1572 GCAAAATATCTGTAGCATGGAGCAGTGGGGCACTTTCACCTCTACCTAGCTTTGGTTGT 1631  
 QY 512 AlaPheThrIleValPheValValLeuTrpValProGluThrIleGlyGly\*\*\*AsnSerArg 531  
 Db 1632 GGATTCACCTGTGGTTTTGTGTGATCTTTGGTCTCTAGACTAAGGAAA-AACGCTTGA 1690  
 QY 532 GlyAspThrIlePheValSerLeu 539  
 Db 1691 AGAGATCCAAAGCTTTGTTCAAGATG 1714

## RESULT 7

AF412060  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Atlg75220/F22H5\_6 mRNA, complete cds.  
 ACCESSION  
 AF412060.1 GI:15724239  
 VERSION  
 FLI CDNA.  
 SOURCE  
 Arabidopsis thaliana (thale cress)  
 ORGANISM

REFERENCE  
 AUTHORS  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1695)  
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banth, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.  
 Arabidopsis cDNA clones  
 Unpublished  
 2 (bases 1 to 1695)  
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banth, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Submitted (21-AUG-2001) Salk Institute Genomic Analysis Laboratory  
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for  
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
 USA  
 Riken Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
 Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H.,  
 Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banth, J., Bowser, L.,  
 Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,  
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,  
 Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,  
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
 Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)  
 contributed equally to this work as PIs.

FEATURES  
 source

1..1695  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="RAPL06-75-L22 (RI3497)"  
 /ecotype="Columbia"  
 1..111  
 12..11475  
 /note="putative integral membrane protein"  
 /codon\_start=1  
 /product="Atlg75220/F22H5\_6"  
 /protein\_id="AAL06513.1"  
 /db\_xref="GI:15724240"  
 /translation="MSPRDDNEARNDLRPFIIHTGSWYRMGSRQSSMMSSQVIRDS  
 SISVLACLIVALGPIQFGFTGYSSTQAAITKDLGTVSEYSEVSGSLNNGVAMVGA  
 IASQIAEYIGRKGLMIAAIPNIIGWLCISFAKDTFLYMGRLLEGFVGIISYVTP  
 VYIAEIAPONMRGGLSGNQLSVITIGIMLAYLLGLFVPRWIRLAVILGILPCTLLIPGLF  
 FIPESPWLAKMGMTDEFSTLQVLRGFDTITVEIKRSVASSTKENTVRFVLDLK  
 RRRYFPLMGIILLVQLQGLINGLVFYSSTIFESAGVTSNAAATFGVCAIQVATA  
 ISTWLVDKAGRRLLITISSVGMTISLIVAAAFYLFKEFVSPDSMYSILSVGVV  
 AMVVFSLGMGPIMLIMSBILPVNIKGLAGSIATLANWFFSWLITMTANLLLAWSGG  
 GTFTYGLVCAFTVVFVTLWVPETKGTLEELQSLFR"  
 1476..1695

## 3' UTR

## ORIGIN

Alignment Scores:  
 Pred No.: 9,49e-121 Length: 1695  
 Score: 1738.50 Matches: 343  
 Percent Similarity: 82.47% Conservative: 71  
 Best Local Similarity: 68.33% Mismatches: 81  
 Query Match: 62.56% Indels: 8  
 DB: 2  
 US-10-051-909-36 (1-553) x AF412060 (1-1695)

QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71  
 Db 33 GAGGAGGCTAGGAATGAT-----TTGCGTCGGCCGTTTATTTCACACCGGAAGC 80  
 QY 72 TrpTyArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91  
 Db 81 TGGTACCGGATGGGTTCGAGACAATCGAGTATG-----ATGGGTTCTCTCAAGTT 131  
 QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111  
 Db 132 ATTCGAGACAGCTCATCTCTGTTCTTGTCATGTTTTCGTTGCTCTTGCCCTATT 191  
 QY 112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu 131  
 Db 192 CAATTCGATCACTTGTGTTTATTCATCTCCAACTCAAGCTGCAATCACCAGGATCTC 251  
 QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151  
 Db 252 GGTTTAACAGTACTGAGTACTCTGTGTTTGGTTCCTTTTCCAATGTGGGTCTATGGTT 311  
 QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyIleGlyArgGlySerLeuMetIle 171  
 Db 312 GGTGCAATTCGCAATGTCAGATTGAGAAATACATTGGACGAAAGGGTCTCTTATGATT 371  
 QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaIleAspAlaSerPhe 191  
 Db 372 GCTGCAATTCCTCAATATCATTTGGATGGCTTTTGCATATCATTTTGCAAAAGATCTCTTTT 431  
 QY 192 LeuTyMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyThrValPro 211  
 Db 432 CTTTACATGGGAAGACTTTTAGAAGGCTTTGGGTTGGGATCATCTCTTACACTGTACCT 491

```
Qy 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 492 GTATATATCGCTGAGATAGCTCCACAAATATGAGAGAGATGGTTCGGTTAAACACG 551
Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251
Db 552 CTTTCTGTGACAAATGGAAATATGTTGGCGTATTACTTGGCCCTCTTGTTCATGAGA 611
Qy 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db 612 ATTCCTGCAGTCTCGGGATATTCGCGTATGATATTTATGATACCAAGCTCTCTTTTCATC 671
Qy 272 ProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGlyThrSerLeu 291
Db 672 CCCGAATCCCTCGCTCGCTGGCAAAATGGGTATGACAGATGAATTTGAAATTCATTA 731
Qy 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db 732 CAACTTCTTCAGGATTTGAGACTGATATTAACCGTTGAGTTAATGAATCAAGAGATCT 791
Qy 312 ValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
Db 792 GTGGCATCATCTCAAAAGCGAAATACAGTTCGGTTTGTAGATCTCAAAAGCAGGAGATAC 851
Qy 332 ArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
Db 852 TATTTCCACATTAATGGTTGGTATAGGTGCTGTACTTCAACAACTTGGTGGAAATTAAT 911
Qy 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
Db 912 GGTGTCTCTATTCTCCAGTACAAATATTTGAATCTCAGAGGTTTACATCAGTAACGCA 971
Qy 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu 391
Db 972 GCAACATTTGGGTGCGGGCTATTTCAGTAGTAGCGACTGCAATATCCATCTGGTGGTG 1031
Qy 392 AspArgAlaGlyArgAlaGlyLeuLeuIleSerThrSerGlyMetThrLeuCysLeu 411
Db 1032 GACAAAGCAGGTCTCGGCTTCTGTTACTATCTCTCGGTGGGATGACGATTAAGCCTT 1091
Qy 412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
Db 1092 GTAAATGTCGAGCTGCTTTCTATCTTAAGGAATTTGTGCTCTCTGATTCAGCATGTAC 1151
Qy 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
Db 1152 AGTTGGCTGACATATTGTCTAGTAGTTGGAGTGTGGCAATGTTGTCTTTTCTCATTTG 1211
Qy 452 GlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleLysSer 471
Db 1212 GGAATGGGACCAATACCGTGGCTCATTTATGTTCTGAGATCCTTCTCTGTAACATAAAGGT 1271
Qy 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
Db 1272 TTAGCTGGAATGATTTGCAACTTAGCCAAATTTGGTTCTTTTCTGGTTGATCACCATGACA 1331
Qy 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer 511
Db 1332 GCAAATTTGCTGTAGCTGAGCAGTGGAGGAATTTCTACTCTGTATGATGATGTTGT 1391
Qy 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArg 531
Db 1392 GCATTCACAGTGTGTTGCTGACTCTATGCTTCTGAGTCTCTGAGACCAAGGCAA-AACTCTTGA 1450
Qy 532 GlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTrpLeuProGluCys 551
Db 1451 AGAACTTCAATCTCTGTTTCAGATGAACAAATTTGAACAACTTCATCTTTGTCCACCTCT 1510
Qy 552 LeuSer 553
Db 1511 CTCTCC 1516
```

RESULT 8

```
AY124845 1464 bp mRNA linear PLN 06-JUL-2002
Arabidopsis thaliana Atlg75220/F22H5_6 mRNA, complete cds.
AY124845
VERSION AY124845.1 GI:21700860
FLI CDNA
Arabisopsis thaliana (thale cress)
Arabisopsis thaliana
Arabisopsis thaliana
REFERENCE
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayaishizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1464)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayaishizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2002) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
COMMENT The two base pair differences do not change the amino acid sequence
of the protein.
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayaishizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGE (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Chan,M.M.,
Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Jones,T.,
Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C.,
Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Davis,R.W.,
Theologis,A., and Ecker,J.R.
Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
Location/Qualifiers
1..1464
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U13497"
/ecotype="Columbia"
/notes="This clone is in pUNI 51"
1..1464
/notes="putative integral membrane protein"
/codon_start=1
/product="Atlg75220/F22H5_6"
/protein_id="AAW70554.1"
/db_xref="GI:21700861"
/translation="MSFRDDNEARNDLRRPFIHTGWSYRMGRSQSMGSSQVIRDS
SISVLACVLIVALGPITQFGTTCGYSSPTQAATIKDLGLTIVSEYVFCSLSNVGMVGA
```

IASGQIARYIGRKSLMIAAIPNIIGWLICISFAKDTSLFYMGRLBGFVGVHISYTP  
 VYIAEAPONMRGGGLSVNQLSVITGIMLAVLGLFVPMRILAVLGLITCLLIPGLF  
 PIPESPRVLMKMGMTDEPESLQVLRGFETDIIVENIKRSVASSTKRNTRVFDLK  
 RRRYFVPMVGLILLVLOQLGGLVLFYSSTIPESAGVTSSNAATFGVGAIOVATA  
 ISTVLWDKAGRLRLITSSVGMGTISLVIAAAYFLKEFVSPDSMDTSMWLSILSVGVV  
 AMVFPISGMGPPIWLMSEILPNIKGLAGSIATLANWFFSWLITWTANLLAWSSG  
 GTFTLYGLVCAFTVTVFVLWVPEIKGKLTLEBLSLFR"

misc\_difference 1463 /note="compared to genome"

/replace="a"

misc\_difference 1464 /note="compared to genome"

/replace="g"

# ORIGIN

## Alignment Scores:

Pred. No.: 1,546-120 Length: 1464  
 Score: 1734.50 Matches: 338  
 Percent Similarity: 83.40% Conservative: 69  
 Best Local Similarity: 69.26% Mismatches: 74  
 Query Match: 62.41% Indels: 8  
 DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x AV124845 (1-1464)

QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71  
 DB 22 GAGGAGGCTAGGAATGAT-----TTGCGTGGCCGTTTATTACACCGGAAGC 69  
 QY 72 TrpTyrArgMetSerSerArgGlnSerPheAlaProGlyThrSerMetAlaVal 91  
 DB 70 TGGTACCGGATGGGTTCCAGACATCGAGTATG-----ATGGGTTCTCTCAAGTT 120  
 QY 92 LeuArgGluSerHisValSerAlaPheLeuGlyCysThrLeuValAlaLeuGlyProIle 111  
 DB 121 ATTCCGAGACAGCTCATCTCTGTTCTTCGATGGTGTTCGATGCTCTTGGCCCTATT 180  
 QY 112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu 131  
 DB 181 CAATTCCGATTCACATTGGTGTATTATCATCTCAAGCTCAAGTCAATCAACAGGATCTC 240  
 QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151  
 DB 241 CGTTTAAACAGTATCTGAGTACTCTGTGTTGGTCTCTTTCCAAATGTGGGTCTATGGTT 300  
 QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGlySerLeuMetIle 171  
 DB 301 GGTGCAATTCAGTGGTTCAGATTGCAGATACATTGGACGAAAGGGTCTCTTATGATT 360  
 QY 172 AlaAlaIleProAsnIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191  
 DB 361 GTTCGCAATTCCTCAATATCATTTGGATGGCTTTGCATATCATTTGCAGAAAGATACTCTTTT 420  
 QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleLeuSerTyrThrValPro 211  
 DB 421 CTTTACATGGGAAGACTTTTAGAGGCTTTGGGTTGGGATCACTCTTACACTGTACT 480  
 QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231  
 DB 481 GTATATATCGTGAGATAGCTCCCAAAATATGAGAGGAGGATTTGGGTTGCGTTAACCA 540  
 QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251  
 DB 541 CTTTCTGTGCAATTTGGAATATGTTGGCGTATTTTACTTTGGCCCTCTTTTCCATGGAGA 600  
 QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271  
 DB 601 ATTTCTTGCAGTTCTGGGATATTGGCGGTATCATTTATGATACACAGGTCTCTTTTTCATC 660  
 QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291  
 DB 661 CCCGATCCCTCGTGGCTGGCCAAATATGGGTATGACAGATGAATTTGAAACTTCAATTA 720

QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311  
 DB 721 CAAGTTCTTCAGAGATTGAGACTGATATTAACCGTTGAAGTTAATGAATCAAGAGATCT 780  
 QY 312 ValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331  
 DB 781 GTGGCATCATCTCAAAACGAAATACAGTTCCGGTTTGTAGATCTCAAAACGAGAGATAC 840  
 QY 332 ArgThrProLeuLeuGlyIleGlyLeuValLeuGlnLeuAsnLeuSerGlyIleAsn 351  
 DB 841 TATTTCCCACTTATGTTGGTATAGGGTTGCTTCTTCAACAACTTGTGGAATTAAT 900  
 QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371  
 DB 901 GGTGTCCTATTCTATTCCAGTACAAATATTGAATCTCGAGGAGTTACATCGAGTAAGCA 960  
 QY 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeu 391  
 DB 961 GCACATTTGGGGTGGGGCTTATTCAGTAGTAGCGACTGCAATATCCATCTGTTGGTG 1020  
 QY 392 AspArgAlaGlyArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeu 411  
 DB 1021 GACAAAGCAGGTCTCGGCTTCTGCTTACTATCTCTTCGGTGGGATGAGATTAGCCTT 1080  
 QY 412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431  
 DB 1081 GTAAATTTGTCAGCTGCTTTCTATCTTAAGGAATTTGTCTCTCTGATTCAGACATGTAC 1140  
 QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451  
 DB 1141 AGTTGGCTGAGCATATTTGTCAGTAGTTGGAGTTGTGCAATGGTGTCTTTTCTCATTTG 1200  
 QY 452 GlyMetGlyAlaIleProTrpLeuMetSerGluIleLeuProValSerIleLysSer 471  
 DB 1201 GGAATGGACCAATACCGTGGCTCATTTATGTCAGATCTCTCTGTGAACATAAAGGGT 1260  
 QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491  
 DB 1261 TTAGCTGGAAGTATTGCAACTCTAGCAATTTGGTCTTTCTTGGTTGATCACCATGACA 1320  
 QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValSer 511  
 DB 1321 GCAAATTTGCTGTAGCTGGAGCAGTGGAGGAATTTCACTCTGTATGAGATGGTTTGT 1380  
 QY 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly\*\*\*AsnSerArg 531  
 DB 1381 GCATTTACAGTGGTGTTCGTGACTCTATGGGTTCTGAGACCAAGGCAA-AACTCTTGA 1439  
 QY 532 GlyAspThrIlePheValSerLeu 539  
 DB 1440 AGAACTTCAATCTTGTTCAGATG 1463

## RESULT 9

AF462803

LOCUS

DEFINITION

AF462803

ACCESSION

AF462803.1

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

AF462803 Arabidopsis thaliana At1g19450/F18014\_20 mRNA linear PLN 08-JAN-2002

AF462803 Arabidopsis thaliana At1g19450/F18014\_20 mRNA sequence.

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

AF462803.1 GI:19087527

Arabidopsis thaliana (thale cress)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 1. (bases 1 to 1914)  
 Cheuk, R., Chen, H., Kim, C.J., Meyers, M.C., Ban, J., Bowser, L.,  
 Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,  
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,  
 Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,  
 Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,  
 Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.  
 and Ecker, J.R.

## Arabidopsis cDNA clones

Unpublished  
2 (bases 1 to 1914)  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Cheuk, R., Chen, H., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L.,  
Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayaishizaki, Y.,  
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,  
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,  
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,  
Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,  
Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.,  
and Ecker, J.R.  
Direct Submission  
Submitted (27-DEC-2001) Salk Institute Genomic Analysis Laboratory  
(SIGNAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN  
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayaishizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the  
sequencing and annotation of the RAPL cDNAs: Cheuk, R., Chen, H.,  
Kim, C.J., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M.,  
Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B.,  
Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S.,  
Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C.,  
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W.,  
Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to  
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)  
contributed equally to this work as PIs.

## FEATURES

Location/Qualifiers  
1..1914  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="RAF09-48-G03 (R21488)"  
/ecotype="Columbia"

## ORIGIN

## Alignment Scores:

Pred. No.: 8.9e-120 Length: 1914  
Score: 1726.50 Matches: 337  
Percent Similarity: 83.03% Conservat: 69  
Best Local Similarity: 68.92% Mismatches: 75  
Query Match: 62.13% Indels: 9  
DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x AF462803 (1-1914)

Qy 52 GluSerGlySerArgHisAspGlyValLeuArgProLeuLeuAenThrGlySer 71  
Db 274 GAAGAGGGAAGAAATGAT-----CTTCGGGCGCGGCTTCTTACACACAGGAAGT 321  
Qy 72 TtpThrArgMetSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91  
Db 322 TGGTATCGGATGGGTTCGACACAAATCTAGTATG-----TTGGAAATCGTCTCAAGTT 372  
Qy 92 LeuArgGluSerHisValSerAlaPheLeuLeuLeuValAlaLeuGlyProIle 111  
Db 373 ATTGAGACAGCTCAATTTCTGCTTAGCTTGTTTGGTTTGGTTCTTGGTCTTAT 432  
Qy 112 GlnPheGlyPheThrSerGlySerSerProThrGlnAspAlaMetValArgAspLeu 131  
Db 433 CAATTCGGATTCACTTGTGGTATTCTCCACCACTCAAGCTGCCATTACTAAGATCTT 492  
Qy 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal 151  
Db 493 GGTTTAACTGATCAGAGTACTCTGTGTTGGGTCTCTATCCAAATGTGGGTGCTATGTT 552

Qy 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGly-ArgGlySerLeuMetIl 171  
Db 553 GGAGCAATTCGCGAGTGGTCAGATTGGTAATACGTTGGACTTGAAAGGGTCTCTGATGAT 612  
Qy 171 eAlaAlaIleProAsnIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPh 191  
Db 613 TGTGCAATTCCTCCAAATATTATTGGATGGCTTTCGATATCATTTGCAAAAGATACTTCTTT 672  
Qy 191 eLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPr 211  
Db 673 TCTTTACATGGGAAGATTGTTAGAGGCTTCGGAGTTGGGATTATCTCTATACGGTTCC 732  
Qy 211 oValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGl 231  
Db 733 CGTATATATAGCTGAGATCGCTCCACAGACCATGAGAGGAGCCTTAGGTTCAAGTAAACA 792  
Qy 231 nLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTzPar 251  
Db 793 GCITTCGTAAACAATTTGGGATATGCTGGGATATTTACTCGGTCTCTTTGTTCCATGGAG 852  
Qy 251 gLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271  
Db 853 AATTCTTCGAGTTTTCGGAGTATTGCCATGTACATTGTTGATACCGGGTCTTTTTTTCAT 912  
Qy 271 eProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLe 291  
Db 913 TCTGTAATCTCTCCGTGGTGGTTCGCAAGATGGGTTTGACAGATGATTTGCAAACTTCATT 972  
Qy 291 uGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAl 311  
Db 973 GCAAGTTCCTCGGGATTTGAGACTGATATTACCGTTGAGGTTTAATGAATCAAGAGATC 1032  
Qy 311 aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysIleTy 331  
Db 1033 TGTGGCATCATCTAGCAACGTTCTGCAGTTCGGTTTGTAGACCTCAAGCGCAGAGATA 1092  
Qy 331 rArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAs 351  
Db 1093 CTATTTCCCACTGATGGTGGTATAGGGCTGCTTGCACCTTCAACAACTTCAGAGAAATCAA 1152  
Qy 351 nGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLe 371  
Db 1153 TGGTGTCTTGTCTATTTCGAGTACAATTTTGAATCTGCGAGGGTGACATCAAGTAGTATG 1212  
Qy 371 uAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLe 391  
Db 1213 GCGCATATTGGAGTTGGCGTTGTTTCAGGTAGTGGCGACTGGGATAGCGACATGGTTGGT 1272  
Qy 391 uAspArgAlaGlyArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLe 411  
Db 1273 GGATAAAGCAGGTCGTGGGCTTCTACTCATGATCTCTTCTATCGGAATGACGATAGCCCT 1332  
Qy 411 uLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTy 431  
Db 1333 CGTGATTTAGCAGTTCGATTTTTTACCTTAAGGAATTTGTATACCTACCTGATTTCAACATGTA 1392  
Qy 431 rTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPh 451  
Db 1393 CAACATTTCTAAGCATGGTTTCCGTAGTTGGAGTTTGGGTATGTTATTTCTTCTCTCTCT 1452  
Qy 451 eGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSe 471  
Db 1453 AGGAATGGGACCAATTTCCATGGCTGATTATGTCGAGATTTCTCCAGTGAATATAAGGG 1512  
Qy 471 rLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetTh 491  
Db 1513 TTTAGCGGGAAGTATAGCAACTTTGTTAAACGTGTTCTGTTAGTGGTACCAATGAC 1572  
Qy 491 rThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyMetValValSe 511  
Db 1573 TGCAAATATGCTGTTAGCATGGAGAGTGGGGGAACCTTTTCACTCTCTACCTTGTGTTG 1632



QY 511 rAlaPheThrIleValPheValValLeuTyrValProGluThrLysGly\*\*\*AsnSerAr 531  
 Db 1633 TCGATTCTACTGGTTTGTGGTCTTTGGTTCCTGAGACTAAAGGAAA-AACGGTTG 1691  
 QY 531 gGlyAspThrIlePheValSerLeu 539  
 Db 1692 AAGAGATCCAACTTGTTCAGATG 1716  
 RESULT 10  
 AX653119 1377 bp DNA linear PAT 22-MAR-2003  
 LOCUS Sequence 2989 from Patent WO03000898.  
 DEFINITION AX653119  
 ACCESSION AX653119  
 VERSION AX653119.1 GI:29155933  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 1 Chang H.S., Chen W., Cooper B., Glazebrook J., Goff S.A., Hou Y.M.,  
 Katagiri F., Quan S., Tao Y., Whitham S., Xie Z., Zhu T. and Zou G.  
 Plant genes involved in defense against pathogens  
 Patent: WO 03000898-A 2989 03-JAN-2003;  
 Syngenta Participations AG (CH)  
 FEATURES  
 source  
 1.1377  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 ORIGIN  
 Alignment Scores: 9.71e-119 Length: 1377  
 Pred. No.: 1710.00 Matches: 358  
 Score: 78.36% Conservative: 44  
 Percent Similarity: 69.79% Mismatches: 39  
 Best Local Similarity: 61.53% Indels: 73  
 Query Match: 6 Gaps: 9  
 DB: 5  
 US-10-051-909-36 (1-553) x AX653119 (1-1377)  
 QY 44 AsnArg--GlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAsp-----Gly 60  
 Db 4 ACACGGGCGGCGGCTCGGCTCGCGATGAGACGCGAAGCGACTACGAGCGCGGC 63  
 QY 61 ValLeuArgArgProLeuLeu--AsnThrGlySerTrpTyrArgMetSerArgGln 79  
 Db 64 GGCATGCGGAAGCCGCTGCTGATGCACACGGGGAGCTGGTACAGATGGGTGCGGCAG 123  
 QY 80 SerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98  
 Db 124 GGGAGCCTCACCGGCGGGGACCTCGTCCATGGCCATCTTACGCGAGTCCACGTCCTC 183  
 QY 99 AlaPheLeuCyThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
 Db 184 GCCTTCCTCTGCACGCTCATCTGCTGCGCTCGGCCCATCTCAGTTCGGATTCCAGCGGGGC 243  
 QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138  
 Db 244 TTCTCTCTCCGACGAGACGCGCATCATCCGAGACCTCGACCTCACCTCTCCGAGTTC 303  
 QY 139 SerAlaPheGlySerLeuSerValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158  
 Db 304 TCGGTGTCGATCGCTGTCACAGCTCGGCGCATGTTGGGCGATGTCAGTGGTCAA 363  
 QY 159 MetAlaGluTyrIleGlyArgLysGly----- 167  
 Db 364 ATGCCCGAGTACATGGCGCGCAAGGGKAGAGATTTTTCCTTTTGGGTACACATC 423  
 QY 168 SerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys 187

Db 424 TCATTGATGATTGCTGCAATTCCAAACATCATTTGGTTGGCTTGCCATCTCTCTTGC AAAA 483  
 QY 188 AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSer 207  
 Db 484 GACTCATCGTTCTTTATATATGAGAGATTGCTCAGGGGTTTGGTGTGGTGTCTCTCT 543  
 QY 208 TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly 227  
 Db 544 TATACGGTGCAGTTTACATAGCAAAATATCAGCTCAAAACATGAGAGGTGCTCTGGC 603  
 QY 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuIleTyrLeuLeuGlyMetPhe 247  
 Db 604 TCAGTGAATCAGTTATCTGTAAACCTGGTGTATATTGTGGCATATTTGCTCGCATGTT 663  
 QY 248 IleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267  
 Db 664 GTTCTCTGGAGGCTCTTCTGCTGTAAATAGGAATCTTGCCTTGCACATGTGTGATACCTGGC 723  
 QY 268 LeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCys 287  
 Db 724 CTATTCTTCAITCCAGAAATCCCAAGATGGTTGGCAAGATGAACATGATGGATGATTTT 783  
 QY 288 GluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAsp 307  
 Db 784 GAGACTTCTTTACAGTTCTCTGAGGGGATTGAGACTGCACATCAGCGCGAAGTGAATGAT 843  
 QY 308 IleLysArgAlaValAlaSerSerSerLysArgThrThrIleSer---PheGlnGluLeu 326  
 Db 844 ATAAAG-----GTAGCATCTTCAAGCAGCAGTATATTTAGCCGATTTGGCTTCGTG 897  
 QY 327 AsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsn 346  
 Db 898 ACACGCAAA----- 906  
 QY 347 LeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyVal 366  
 Db 907 -----ATTTCTTCAGGTC 921  
 QY 367 ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyVal 386  
 Db 922 ACAAACAGTGCATTTGGCTACATGTGCATTTGGTGTCTATCCAG----- 963  
 QY 387 ThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuIleSerThrSerGly 406  
 Db 964 -----ATCTCTTCTGCTGGG 978  
 QY 407 MetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGln 426  
 Db 979 ATGACTCTAAGCCTCTTGCAGTTGCTGTATTTTCTCAAGGATAGCATTTCAAA 1038  
 QY 427 AspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheVal 446  
 Db 1039 GATTTCTACATGTACTACACCTTAAGTATGATCTCTTGGTGTCTTGTGGCTTTGTA 1098  
 QY 447 IleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuPro 466  
 Db 1099 ATTCGCTTCTCTCTGCGTATGGGTGCCATTCATGGATCATATGTAGAGATCTCTCCG 1158  
 QY 467 ValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPhe 486  
 Db 1159 GTTAGTATCAAGAGTCTCCGAGGAAGCTTTCGACGCTCGCAACTGGCTTACATCTTT 1218  
 QY 487 AlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSer 506  
 Db 1219 GGAATAACAATGACAGCAAACTTGATGCTTAGCTGGAGTGTCTGGAGGACCTTTGTGTCC 1278  
 QY 507 TyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLys 526  
 Db 1279 TACATGGTCTGTGAGTCTTTCACTCTGTTGCTCATCTTTGGTGGTCCAGACAAAA 1338  
 QY 527 Gly\*\*\*AsnSerArgGlyAspThrIlePhePheSerLeu 539  
 Db 1339 GGAAG-AACTCTCGAAGAGATACATGCTCTCCGCTG 1376



```

RESULT 11
BT014234      1825 bp      mRNA      linear      PLN 11-MAY-2004
LOCUS      Lycopersicon esculentum clone 133433R, mRNA sequence.
DEFINITION
ACCESSION      BT014234
VERSION
KEYWORDS      FLI_CDNA.
SOURCE
ORGANISM      Lycopersicon esculentum (tomato)
OCUT      Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1825)
REFERENCE
AUTHORS      Kirkness,E.F., Wang,W. and Vazelle,A.
TITLE      Direct Submission
JOURNAL      Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1..1825
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="133433R"
/tissue_type="fruit"
/notes="TXHM68"

ORIGIN
Alignment Scores:
Pred. No.:      2,7e-117      Length:      1825
Score:      1893.00      Matches:      328
Percent Similarity:      82.85%      Conservative:      73
Best Local Similarity:      67.77%      Mismatches:      79
Query Match:      60.92%      Indels:      5
DB:      8      Gaps:      2

US-10-051-909-36 (1-553) x BT014234 (1-1825)

Qy      57      AspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSer      76
Db      122      AACATGAAGGGATCTGAAGAACCCATTTTACATACCTGGGAGTTGGTATCGTATGGGG      181
Qy      77      SerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHis      96
Db      182      TCAAGACACTAGTAGT-----ATGGGTCTCTCTCAAGCTATTCTGTATAGTTCT      232
Qy      97      ValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProIleGlnPheGlyPheThr      116
Db      233      ATTCTGTCTTGTCTGTCTTCATAGTGTCTTGGGCCCAATCCAGTTTGGTTTCACT      292
Qy      117      SerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSer      136
Db      293      TGTGGCTACACTTCACCAAGCAAAATTGCAATTACCAAGAGCTGAAGCTCACTGTTCT      352
Qy      137      GluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSer      156
Db      353      GAGTTCTCTTGTTCGGTCTTTATCCAAATGTGGGTGCTATGGTTGGTGCATATCTAGT      412
Qy      157      GlyGlnMetAlaGluThrIleGlyArgGlySerLeuMetIleAlaIleProAsn      176
Db      413      GGTCAAAATCTGAGTACATGGGCCGAAGGGTCTTTAATGATAGCTGCTATACCAAT      472
Qy      177      IleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArg      196
Db      473      ATCATTTGGTTGGCTTCCCATCTCATTTGCCAAAGATCCTTTCTTATACATGGGAAGA      532
Qy      197      LeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGlu      216
Db      533      TTGTTGGAGGGCTTGGTGGGAATAATCTCTTATGTGTTCTCTGTATATATTGTGAG      592
Qy      217      IleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPhe      236
Db      593      ATTGCACCTCAGAATAATGAGAGGGGTCTAGGGTCTGTAAACAGCTCTCTGTACAAAT      652

```

## RESULT 12

AX653120

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AX653120

Sequence 2990 from Patent WO03000898.

AX653120

AX653120.1

GI:29155934

Oryza sativa

1428 bp

DNA

linear

PAT 22-MAR-2003

ORGANISM *Oryza sativa*  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  
 AUTHORS Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,  
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
 TITLE Plant genes involved in defense against pathogens  
 JOURNAL Patent: WO 0300898-A 2990 03-JAN-2003;  
 Syngenta Participations AG (CH)

FEATURES  
 Location/Qualifiers  
 source 1..1428  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4330"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,43e-109 Length: 1428  
 Score: 1577.00 Matches: 318  
 Percent Similarity: 85.8% Conservative: 34  
 Best Local Similarity: 77.56% Mismatches: 32  
 Query Match: 56.75% Indels: 27  
 DB: 6 Gaps: 2

US-10-051-909-36 (1-553) x AX653120 (1-1428)

QY 130 AspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGly 149  
 DB 277 GATCTGTATACAAACAGCATATGTTCTCGGCTTTTCGGTTTCGCTGCCAACGTCGGCGC 336  
 QY 150 MetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGlySerLeu 169  
 DB 337 ATGGTCGGAGCGATCGCCAGTGGCAGATGGCGAGTACATTGGCCGGAAAGGGTGTGTG 396  
 QY 170 MetIleAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaIleAspAla 189  
 DB 397 ATAATGGCGCGTTCCTAAACATCATTTGTTGGCTGCCATCTCTTGCACAAAGACGGC 456  
 QY 190 SerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209  
 DB 457 TCATTTTATACATGGGACGCTGCTTGAAGGTTTGGTGTGGTGTATATATATATACG 516  
 QY 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229  
 DB 517 GTGCCAGTATACATGACAGATATCTCATCAGAACACAGAGAGACCTTGGCTCCGTG 576  
 QY 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePro 249  
 DB 577 AACCAATGTCCTTACCATTTGGTATCTTGTGGCTTATTTGTAGGCATGTTGTCTCT 636  
 QY 250 TrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhe 269  
 DB 637 TGGAGGCTGCTTGCAGTGATAGGAAGCATCCCATGTACATTTGTAATACCTGCTATTTC 696  
 QY 270 PheIleProGluSerProArgTyrTrpLeuAlaIleMetAsnLeuThrGluAspCysGluThr 289  
 DB 697 TTCATTCGGAAATCCCAAGATGCTGGCAAAATGAATGATGATGATTTGAGGCT 756  
 QY 290 SerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLys 309  
 DB 757 TCTCTACAAGTTTGGAGGGGTTTGGAGCTGATATATTACTGCAGAA 801  
 QY 310 ArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys 329  
 DB 802 AGACAGTAGCATCGGCAAAACAAAGAACACACAGTCGGTTTAAAGAGTTGAACCAAAAG 861  
 QY 330 LysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGly 349  
 DB 862 AAATACCGCACTCCCTTACTAGTAGGAATCGGCTTCTGTACTTCAGAAATCTAAGTGA 921  
 QY 350 IleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaIleValThrAsnSer 369

DB 922 ATAAATGGTATTCTGTTTATGCAAGTAGAATCTTCAGAGATCGCGGTTTACAAACAGT 981  
 QY 370 AspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrp 389  
 DB 982 GACTTGGCCATCTGCATAGGAGCAATTCAG 1014  
 QY 390 LeuLeuAspAlaGlyArgIleLeuLeuIleIleSerThrSerGlyMetThrLeu 409  
 DB 1015 -----ATCTTACTGCTGGATGACTCTA 1038  
 QY 410 CysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsn 429  
 DB 1039 AGCCTTCTTGCATATCTGTTGATATTTCTGGAGGTATATTTTCATGATTTCTCAI 1098  
 QY 430 SerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPhe 449  
 DB 1099 TCGTTTCTACATCTTAAGTATGATCTCTCTGGTGTCTTCTGTGGCTTATATCATCACCTTT 1158  
 QY 450 SerPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIle 469  
 DB 1159 TCCTTGGCATGGGTGCCATTCATGGGTGCATGATGCTGAGATCCTCCAGTTAGCATC 1218  
 QY 470 LysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThr 489  
 DB 1219 AAGAGTCTTGGGGAAGCTTTGGCACACTTGCACACATGCTTACATCTCTGGGCAATAACA 1278  
 QY 490 MetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetVal 509  
 DB 1279 ATGACAGCAATTTGTTACTCAGCTGGAGTGCTGGAGGAACATTTTGTCTTACATGATT 1338  
 QY 510 ValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly\*\*\*Aan 529  
 DB 1339 GTGAGTCTTCTCACTTCGTGTTCTCATATTTTGGTGCCAGACAAAGGGAG-AAC 1397  
 QY 530 SerArgGlyAspThrIlePheValSerLeu 539  
 DB 1398 TCTGGAGGAGATACAGTTTTCGTTCTGCTA 1427

RESULT 13  
 AC120986 204649 bp DNA linear PLN 02-SEP-2004  
 LOCUS Oryza sativa (japonica cultivar-group) chromosome 5 clone  
 DEFINITION OJ1781\_H11, complete sequence.  
 AC120986  
 VERSION AC120986.2 GI:46518554  
 KEYWORDS HTG.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 REFERENCE 1 (bases 1 to 204649)  
 AUTHORS Chow, T.-Y., Hsing, Y.-I. C., Chen, C.-S., Chen, H.-H., Liu, S.-M.,  
 Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R.,  
 Chen, Y.-J., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H.,  
 Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kuo, P.-I., Lee, M.-C.,  
 Liu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y.,  
 Yu, S.-W., Wu, H.-P. and Shaw, J.-F.  
 TITLE Oryza sativa BAC OJ1781\_H11 genomic sequence  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 204649)  
 AUTHORS Chow, T.-Y. and Hsing, Y.-I. C.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-MAY-2002) Institute of Botany, Academia Sinica, 128,  
 Section 2, Academia Road, Nankang, Taipei 11529, Taiwan  
 REFERENCE 3 (bases 1 to 204649)  
 AUTHORS Hsing, Y.-I. C. and Chow, T.-Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-MAR-2003) Institute of Botany, Academia Sinica, 128,  
 Section 2, Yien-chu-yuan Road, Nankang, Taipei 11529, Taiwan  
 REFERENCE 4 (bases 1 to 204649)  
 AUTHORS Chow, T.-Y.  
 TITLE Direct Submission

## JOURNAL

REFERENCE  
AUTHORS

## TITLE

## JOURNAL

## COMMENT

Submitted (23-APR-2004) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan  
 5 (bases 1 to 204649)  
 Chow.T.-Y. and Hsing,Y.-I.C.  
 Direct Submission  
 Submitted (02-SEP-2004) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan  
 On Apr 23, 2004 this sequence version replaced gi:20806177.  
 The nucleotide sequence of this BAC clone was generated by combining Monsanto and ASPGC-Taiwan sequencing data. Genes were predicted from the integrated results of the following: BLASTN2.0, BLASTX2.0, GENSCAN (Chris Burge,  
<http://genes.mit.edu/GENSCAN.html>), Fgenesh  
<http://www.softberry.com/>), GlimmerK  
<http://www.tigr.org/sofclab/glimmer.html>), TWINSKAN  
<http://genes.cs.wustl.edu/>) and GeneSplicer  
<http://www.tigr.org/cdb/Genesplicer/index.shtml>). The sequence was searched against the Swiss-Prot-TrEMBL protein database, the NCBI Plant EST database, the TIGR Rice Gene Index and the rice full-length cDNA database (KOME,  
<http://cdna01.dna.affrc.go.jp/cDNA/>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. This clone overlaps with OSJNB0053D02 (accession # AC124143) and OSJNB0040806 (accession # AC120990).

FEATURES  
source

Location/Qualifiers  
 1..204649  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="5"  
 /clone="OJ1781\_H11"  
 14532..17456  
 /gene="OJ1781\_H11.1"  
 /notes="WRKY transcription factor"  
 join(14532..15049,16339..16482,16859..17456)  
 /gene="OJ1781\_H11.1"  
 /codon\_start=1  
 /product="WRKY transcription factor"  
 /protein\_id="AAU10654.1"  
 /db\_xref="GI:51854273"  
 YHHMQESATTPPLDYEPAGFDDVAPLEVEVKRELIVDVGVLPPGGGASAAAA  
 AAVAGMPTNMSVSTSSACGGVGGAGGDEESAGCKKBEEDGGDDDDGKGGSTP  
 KGDGDEDNKNGKGGKGGKGRQPRFAPMTKSEVDHLEDGYRWRKYQKQAKNSP  
 FPRSYRCTTQKCPVKRVERSYQDAAVITTYEGKHTHPIPATLRGTAHLGAAAAA  
 HHHGGLYHHPHGFHFAAVGHRLLPQPHDALGGGLLAPPHAQHLHAMQHOLAAAAA  
 SGSLHRAAMQMPQDPHAGLVALIASTGASTTPPPPPATGSAATTLRQHFMA  
 QDYGLLQDMFIPSPFLHNDANNR"  
 complement(24593..28735)  
 /gene="OJ1781\_H11.2"  
 /notes="polyprotein"  
 complement(join(24593..25737,25868..26267,26420..28735))  
 /gene="OJ1781\_H11.2"  
 /codon\_start=1  
 /product="polyprotein"  
 /protein\_id="AAU10655.1"  
 /db\_xref="GI:51854274"  
 LPAMALRYVVDVAFILSKLKYTSSTVTRTHIANVMFIVLFWIKICRIITKPP  
 TIQKLRSLVFAASLKPFAHFDGSKKAKARALLMLVMQCFVSRGKSPPEPLF  
 PEEAKFEATDCLFRGALISVLADISIVDVMHMPGKMMWDALEAKFPGSDVDSLEY  
 MEQFYDKYVDDCSVQAQHNQAKELNNNCELPDKFVAGIITAKLPSPWTDFAI  
 SLKHROEFVFTDLIGSLVBEKRAQDNRGKIKTEGGSANLVQKNKPNASHNNKKV  
 KPDVKPATTNFKKKGKAKGDCFCVCGKPHWAKCPCPKDKKANWVISEGGTIG  
 YGRFLPTVLSVCHSLDHWDTGANTIHVFDIISLPSSVQVGRGSLMNGSLVAVHGV  
 YGVDLKFTSKIVOLKNQHVPSIKNMLVSGSLLCRDGPRLVFESNKCVVSKYGTFIG  
 KYDSDGGLFRSLDDMCNKVNVHVSDDDDSNVWHSQCHVNFQCMTRLANNSLIKPF

## gene

## CDS

## gene

## CDS

## gene

## CDS

## gene

## CDS

## gene

## mRNA

## mRNA

## gene

## CDS

TLVKGKCHTCVQTKQPKHKSSSEARKLAPELIVHSNLCENMGVLTGKGKGYFMTLI  
 DDCTRFCYVLLKTDQEVULHYPKIYKAEVFNQLEKIKRLSDRGGEYSPNETFSCK  
 EYGIHERTHSYSPQSGVAERKNRLLTEVMNMLDTAGLSKEMWGAELITACHVUNR  
 IPTKKEVTSFEERKKLNSHLRTWGLAKRCHTFENEFPMKNAPSTSSOBPLISP  
 EHFVLIETHDQTEENSEENNVATRKSKQRQRTAKSGFDDYIVYLVDDTPRTIEEAYS  
 SPVDYKFALESEEMDSIMSGTWIECPYCGKPVGCKWPKKLPDPDASHGLLVHQ  
 MDVTAFLNGELEEEIYMDPGYVLEGOEGMVCKLLKSLYGLQKAPKQWYKXFDNVL  
 TSTGFVNEADKCVYRNGGGGVLICLVNDILIFGTSINVTIEVKDELSEKPKKD  
 LGVADVILNKLRLGDEGRITLVQSHVDKLSRFGYSCKAAFTYDFSVLURKRR  
 IARDQRYQSIITGLMYLASATRPDISFAVSKSRFVSNPDDHMHALERSVLRKGT  
 MSYGHVAGYKPVKEGYSNDSIDAETKATSGYVFTLGGGAVSKCKQIILTRST  
 MEETLHATATVAEAWRELLMDPVVEKPIPAILLMNCNDQTVIIKVNSSKDNMKSS  
 RHVKRSLKSVRKQKNSRVIALDVV"  
 37831..40988  
 /gene="OJ1781\_H11.3"  
 /notes="unknown protein"  
 join(37831..38026,38128..38219,39503..39584,40203..40619,  
 40704..40849,40965..40988)  
 /gene="OJ1781\_H11.3"  
 /notes="similar to rice EST BX901392, AK099836"  
 /codon\_start=1  
 /product="unknown protein"  
 /protein\_id="AAU10656.1"  
 /db\_xref="GI:51854275"  
 /translation="MLTLMKNCNDDDIGSEFYDSFNLNDNTIENVHSONNKLAKT  
 GFNQYETGFSLPCGDQYDYLRTAQDNKSGEGADPMDDGTFTLLKDLVTES  
 RTSSPYEQKHQNLNLYNQDGHGPTTTNPTWDLTQSCNLHSLTSDSPVFNQV  
 SAGVFNATSNDFIDELATISGRYVLPSEIKTEEGSIAIDANSPATATMVCNS  
 VHFHWIDQNLGTPDLADLPDIYDPSFLTPRKNTILVDLDTLIHSSAVDRDGA  
 DFPFPMYHGLKEHTVYVKKRPHVDTPQLQVSEFMKVAILRI"  
 43029..45134  
 /gene="OJ1781\_H11.4"  
 /notes="unknown protein"  
 join(43029..43349,43462..43596,44124..44234,44787..45134)  
 /gene="OJ1781\_H11.4"  
 /notes="similar to rice EST AK099641"  
 /codon\_start=1  
 /product="unknown protein"  
 /protein\_id="AAU10657.1"  
 /db\_xref="GI:51854276"  
 /translation="MAPPATASASAPVSLFLSLPSSRCGLPAPHTLPPRRLL  
 ALAPRLGAALLSSLDQAEEREYDDEEBELVVGVYSGAHGVRGDVLVSPRTDPQ  
 LRPAFGKRLARAGAKQKVEFELVRGKRAHTKKSMTVDPGVSDVEARQIVGSA  
 ILVKGDRPMBEDEIYSLDLVGMRVIVKDTGKLVGTQVQVFNFGAGDLLQVWVGTE  
 DTVSQNSQGSSTGSEHVIWIPFAEDIVPIDMESREMWITPPKGLLELNAKSDKRS  
 KERRAMVRSASSFCMLCSMYKAFRT"  
 complement(49191..53735)  
 /gene="OJ1781\_H11.5"  
 /notes="putative mitogen-activated protein kinase"  
 complement(join(49191..49621,49723..49805,50006..50098,  
 50177..50388,50518..50673,50769..50918,51360..51419,  
 51534..51661,52173..52581,53679..53735))  
 /gene="OJ1781\_H11.5"  
 /product="putative mitogen-activated protein kinase"  
 /notes="similar to rice EST CF306885, AK099472, AK061645"  
 complement(join(49191..49621,49723..49805,50006..50098,  
 50177..50316))  
 /gene="OJ1781\_H11.5"  
 /product="putative mitogen-activated protein kinase"  
 /notes="similar to rice EST AK100389"  
 complement(62202..63959)  
 /gene="OJ1781\_H11.6"  
 /notes="putative eukaryotic translation initiation factor  
 (eIF3d)"  
 complement(62202..63959)  
 /gene="OJ1781\_H11.6"  
 /notes="similar to rice EST AK104093"  
 /codon\_start=1  
 /product="putative eukaryotic translation initiation  
 factor (eIF3d)"  
 /protein\_id="AAU10658.1"  
 /db\_xref="GI:51854277"  
 /translation="MGFDVGVAFNPEGWPPDTPAPASLGGGAAAASIPPAFFSRS

DKLGRIADWTRNPAGPAALAAASRDSVDFDFTSVDSDLSAAAEDSSFLVDKAPPPRHP  
 RGPKNRNQRLPQRDERVEARERARRERHFNKNSHHHPFRGNQSSS  
 AKSVDIQDWTMRREQIPFANFTKLFSFSAQDPEDLLLCGAVEYDRAFRVNPKAAR  
 RLERPKSNFVKVTFDDPVRLRLAEDKATVFATDAI LALMCTPRSHSDI VVOR  
 VGNKLFDFKRGSDQLDLSVNETAQEQLPENKNDINGASHLAVEATYINQNFQQVLL  
 RDGEKTFDEPNPFASGEESAAGYRYRWKLDDETSI VARCEVHAVNADPGGGRQF  
 LTLNALNFDFPKITGVDRWQKLETQRGAVLATELKNANKLARKWTQCALLAGAMMKL  
 GYVSRVPRDHYNHAILTVMGYKPRDFATQINLNTSNMGI VSI VDI CMKPEEGKYV  
 LVKDPAKQVRIYEVPSDFENDYVEEPLPEEQVRPPSDVDVDATEMDAAAAEABEN  
 NAAAGAGGEGKSAEAAAA"  
 complement (65704, .66784)  
 /gene="Q11781.H11.7"  
 /note="unknown protein"  
 CDS  
 complement (join(65704, .65774, 65857, .65914, 66247, .66306,  
 66414, .66461, 66620, .66682, 66764, .66784))  
 /gene="Q11781.H11.7"  
 /note="similar to rice EST AK058475"  
 /codon\_start=1  
 /product="unknown protein"  
 /protein\_id="AAU10659.1"  
 /db\_xref="GI:51854278"

Alignment Scores:  
 Pred. No.: 1.68e-96 Length: 204649  
 Score: 1453.50 Matches: 426  
 Percent Similarity: 37.12% Conservative: 45  
 Best Local Similarity: 33.57% Mismatches: 54  
 Query Match: 52.30% Indels: 747  
 DB: 8 Gaps: 19

US-10-051-909-36 (1-553) x AC120986 (1-204649)  
 QY 3 SerSerSerPheArgProAlaGlyLysLysLysLysLysAanGlnGlyLeuArg 22  
 Db 118274 TCTCTCTCATCATCTCTCCGTTGAGCGGCGCGC----- 118309  
 QY 23 ArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGly 42  
 Db 118310 -----GGCGAGCCGAGCGAGCGAGCGAGCGCGCGCGATGAACGCGCGCGG 118357  
 QY 43 SerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAsp----- 59  
 Db 118358 AGCAGCGGGAGGAGGCGTGTGCGCCGACGAGAGCGGCGGACGACGAGGCGGCGGCGC 118417  
 QY 60 ---GlyValLeuArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArg 78  
 Db 118418 GCGCGCGCTTTCGGGAACCGCTGCTGAACACGCGGAGCTGGTACAGGATGGGATCGCGG 118477  
 QY 79 GlnSerSerPheAlaProGlyThr-SerSerMetAlaValLeuArgGluSerHisValSer 98  
 Db 118478 TCCAGCCTCGCGCC-----TCTTCCATGGCGCGCATCCGGAGTCCCACGCTCC 118528  
 QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
 Db 118529 GCCTTCTCTGCAGCTCATCTGTCGCGCTCGGCCCCATCCCAATTCGGATTCCACGCGGC 118588  
 QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu--- 137  
 Db 118589 TTTCTCTCACCACCCAGGAGCGCATCATCCGCGACCTCAAGCTCTCCATCTCCGAGTTC 118648  
 QY 137 ----- 137  
 Db 118649 CCGTCCCGTTCCAGCTCGGACNTTGGCGTGTCTTTTGTGCTTCTCTGTCTTATTCGGTT 118708  
 QY 138 -----PheSerAlaPheGlySerLeuSerAs 146  
 Db 118709 TCTTGTAGTGTGAAGTTTATCTTTCTGTATGATGACGTTCTCGGCTTTTCGGTTGCTGCCAA 118768  
 QY 146 nValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTrpIleGlyArgGly 166  
 Db 118769 CGTGCGGCCATGTGTGGAGGATCGCCAGTCGGGAGATGCGGAGTACATTGGCCGGAA 118828  
 QY 166 sGly----- 167

||||  
 Db 118829 AGGGGTGAGAAATTTGTGCTTCATTGTGTCCTATCTGTCTTCCACATAGAGGCCATCCT 118889  
 QY 167 ----- 167  
 Db 118889 GTATCTCGACGACGAACAATTTGATACCTTTGTAGTGTGACAAGGTCAATACGTCACTGTTG 118948  
 QY 167 ----- 167  
 Db 118949 TATGATTTCCGTAACACCAATAATTTGCCATTAAGAAAGATGAGTGGGATTAGAAATGC 119008  
 QY 167 ----- 167  
 Db 119009 TGGTCATTTCTGTGNACTGTAGCTTGTGAAATATGAAACCAAAATTTTCAGTTCAGT 119068  
 QY 167 ----- 167  
 Db 119069 TAGCTATACCTATTGGGGTTTGACCTTAATTTCAAGTTGACGAGTCTTTAGTTGTGATT 119128  
 QY 167 ----- 167  
 Db 119129 GCCATTGATCTGGCAAAAGAGAGTTAGTAGACAGAAATGTGTAGATGAAGTTTTTGT 119188  
 QY 167 ----- 167  
 Db 119189 GACATTTTCTTTGGTTTCATCTTCAATCAATCACTTAATCAGATCAGTAGTTAACTACAGTG 119248  
 QY 168 -----SerLeuMetIleAl 172  
 Db 119249 AAGCTCTATTATTTTACTCTAATGGTCTGTGCTGAATCTCCAGTCGTTGATAATTGC 119308  
 QY 172 aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys----- 187  
 Db 119309 GCGGTTCTTAACATCATTTGGTTGCTTGCATCTCTTTCGCAAAAGTACGTGTACATA 119368  
 QY 187 ----- 187  
 Db 119369 TTTCTCAAGCTTATTTTCTCTGTTCTCTTTTCTACTAGTTTCTATTCTTGGCGATCTT 119428  
 QY 188 -----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGI 204  
 Db 119429 TGTTTTTCAGACCGGTCATTTTATACATGGGACGCTCTCTTGAAGGTTTGTGTGTGG 119488  
 QY 204 YIleIleSerTyrThr----- 209  
 Db 119489 TGTATATCATATACGCTCTGTATTTTACCAACTTCATTTTCATGTTTATTTGCTGTAGA 119548  
 QY 209 ----- 209  
 Db 119549 ATGTTCTGTGCTTCACTGCTTGAGTACTTATCTGACTGAGACTGAGATGCTGTAGAAT 119608  
 QY 209 ----- 209  
 Db 119609 GAGCGAGGACATCCTTTGCTAATGAAGCTTTTATACATTTTGTGTCGAGAACTTTTG 119668  
 QY 210 -----Va 210  
 Db 119669 GATCAGTTAATAGTTAGTGCATGTATAAACCTTTAATGTGTGAATTATATCGTACCAGT 119728  
 QY 210 lproValTyrIleAlaGluIleSerProGlnAsnMetArgGlyValAlaLeuGlySerValAs 230  
 Db 119729 GCCAGTATACATAGCAGATATCTCATCAGAACACAAAGAGGAGCACTTGGCTCCGTGAA 119788  
 QY 230 nGln----- 231  
 Db 119789 CCAGGTGTGCTTATTCGTCAGAAAGCATCTTATCTATTCGCGAGGAAATTTATGTTGTTT 119848  
 QY 232 -----LeuSerValThrPheGlyIlePheL 240  
 Db 119849 ATGTTGGACTTTACCATGCAGTCTTTTCATTTATGTAGTTGTTCGTTACCATTTGGTATCTGT 119908  
 QY 240 euAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIle----- 256

```
Db 119909 TGGCCTATTTGCTAGGCATGTTTGTCTTGGAGGCTGCTTGCAGTGATAGGTACAAATC 119968
Qy 256 -----
Db 119969 AATACCTGATGTCGAATGTTTCTATTAGGATATGTAATAATAATGAATTTT 120028
Qy 256 -----
Db 120029 AACTTTTGATATTGCTAGTACCTTTTGGACTTATTGCAATTTGTATGTCGTTAAATGT 120088
Qy 257 -----GlyAlaLeuProC 261
Db 120089 CTGGTACCTCGATTCTGATTCAAAATATCTGGGTGCCATGCTTGTAGGAAGCATCCCAT 120148
Qy 261 YstMetLeuLeuProGlyLeuPhePheLeuProGluSerProArgTIPLeu----- 278
Db 120149 GTACATTGTTAAATACCTGGTCTATTCTTCATTCGGGAATCCCAAGATGCGCTGTATGTT 120208
Qy 278 -----
Db 120209 ATACACCTAAGTTGGATATTGTATCACAATCTGAAATCTGTAGCAAGCACTTTCCTTACA 120268
Qy 279 -----AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuA 295
Db 120269 TCTATATGCAAGGCAAAATGAATAATGATGATGATTTTGAAGCTTCTCTACAAATTTTGA 120328
Qy 295 rgGlyPheGluThrAspIleThrThrGluValAsnAspIleLys----- 309
Db 120329 GGGGTTTGGAGCTGATATTACTGCAGAAGTGAATGATATAAAGGCAAGCAATTTTATT 120388
Qy 309 -----
Db 120389 TTCTTGATTAACTGACTACAAGTTATTTTGCGGTTTTCGCGTATTTCATCTTAGTAAGATT 120448
Qy 309 -----
Db 120449 TTCACTAACACAACACTTGGGAGATATGTACTTTCATTTTTCGGAAGACATAATATGGTT 120508
Qy 310 -----ArgAlaValAlaSerSerSerLysArgThrThrIleSerPhe 323
Db 120509 CAAATCCCTGTTATGCAGAGAGCAGTAGCATCGGCAAAACAAGAACCCACAGTCCGTTTT 120568
Qy 324 GlnGluLeuGlnLysLysThrArgThrProLeuLeuLeu----- 337
Db 120569 AAAGAGTTGAACCAAAAGAAATACCGCACTCCCTTACTGCT- AACTTATACATACCCCTTTT 120627
Qy 337 -----
Db 120628 GTTCAGTTTATACCAAGTCATTAAATCATGTTTGTGTAATGGTAACACTACGCTTTATGTTGA 120687
Qy 338 -----GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
Db 120688 CAGATAGGAAGTGGCTTCTTGTACTTTCAGAACTCTAAGTGGATAAATGTAATGTAATCTGTTT 120747
Qy 356 TyrAlaSerSerIlePheLysAlaAlaGly----- 365
Db 120748 TATGCAAGTAGAATCTTTCAGAGATGCAGGTAAGTTGTGGATATCTCTTATAGTGCCTTAG 120807
Qy 365 -----
Db 120808 TACTACAGTCAGCTGTAGTAAATTTTGGCTGTGTAATCTTAAACGCAAGAACTTCAGG 120867
Qy 366 Val-ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln----- 380
Db 120868 GTTTACAAACAGTGACTTGGCCACATGTGCACTAGGAGCAATTCAGGTAAGTATAGGA 120927
Qy 380 -----
Db 120928 ATATCTCTGATCTCATCTCTGTAACTTATATCAAGTTGCTTACAGAAATCGTTGGATTACA 120987
Qy 381 -----ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLe 399
Db 120988 TTAGGTTCTTGCCACTGGAGTTACAAATGTTTACTTGCAGACAGCTGGTTCGACGGATGCT 121047
```

## RESULT 14

AC025808/c

LOCUS

DEFINITION

AC025808 120977 bp DNA linear PLN 11-OCT-2000  
Genomic sequence for Arabidopsis thaliana BAC F18014 from

ACCESSION	chromosome I, complete sequence.	
VERSION	AC025808	
KEYWORDS	AC025808.8 GI:7636235	
SOURCE	HTG.	
ORGANISM	Arabidopsis thaliana (thale cress)	
REFERENCE	Arabidopsis thaliana	
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 120977)	
TITLE	Shinn, P., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, Q., Chiu, C., Chioi, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Federspiel, N.A., Theologis, A. and Ecker, J.R. Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome I	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 120977)	
AUTHORS	Ecker, J.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA	
REFERENCE	3 (bases 1 to 120977)	
AUTHORS	Ecker, J.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (22-APR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA	
REFERENCE	4 (bases 1 to 120977)	
AUTHORS	Cheuk, R., Shinn, P., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, B., Chin, C., Chioi, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA	
REFERENCE	5 (bases 1 to 120977)	
AUTHORS	Chao, Q., Brooks, S., Buehler, E., Johnson-Hopson, C., Khan, S., Kim, C., Shinn, P., Altafi, H., Bei, B., Chin, C., Chioi, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, S., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA	
COMMENT	On Apr 22, 2000 this sequence version replaced gi:7543634.	
FEATURES	Location/Qualifiers	
source	1..120977	
	/organism="Arabidopsis thaliana"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:3702"	
	/chromosome="1"	
	/clone="F18014"	
	90..2159	
	/note="hypothetical protein"	
	/codon_start=1	
	/evidence=not experimental	
	/product="F18014.1"	
	/protein_id="AAF79419.1"	
	/db_xref="GI:8778411"	
	/translation="MISFEVSPDVFTCSIVVNAVCRSGNVDKAMVFAKETSSLGLEL NVVTVNSLINGAMIGDEGTRVLRILMSRGRVNVTVTSLIKGYCKKGLMEAEH	
CDS		
	VFELLKEKLVADQHMVGVLMDGYCRTGQIRDAVRVHDNMIEIGVTRNTTNCNSLING YKSGQLVEARQIFSRMNDVSLKPDHVTYNTLVGDCRAGYVDEALKCLCDMQCQKEVV PTVMYNTILLKGYSRIGAFHDVLSLWMKMLKRGVNADEISCTLLALFKLGDNFNAH KLWENVLARGLLTDTILNVMISGLCKMEKVNKEILDNVNI FRCKPAVQTVQALSH GYKVNGLKEAFAYKEYMERKGIPTTIEMTYNTLISGAFKYRHLNKVADVLEIRLQIDA CLLQKTVDFDLLPGYOSLKEFLEASATTCLKTOKIAESVENSTPKLLIVPNNIVVN VAIGLCKAGLEDARKLFSDLLSDFRIFDEYTYTILIHGCAIAGDINKAFTLRDEM ALKGI PNVTYNALIKGLCKGNVDRQARLLHKLPKGTTPNAITNTLTIDGLVKSQ NVASARMLKEKMIKGLVRGSKQGDVDIPKEVVLDPKVGSTGVIEMNSNELYDVR RVSEAVV"	
	complement(4647..5702)	
	/note="putative glycosyltransferase-like protein	
	emb CA542305.1; similar to EST gb A1998490.1"	
	/codon_start=1	
	/evidence=not experimental	
	/product="F18014.2"	
	/protein_id="AAF79456.1"	
	/db_xref="GI:8778448"	
	/translation="MSQHLLLLILLSSLLHLPISATTIIQKFKPAQFVNSADCPLI DSESDDVVAKPIFCRRRAVHVAMTLDAAVIRGSVAVLISVLOHSCPENIVFHVVA SASADASSLRATISSPFDFTVYFVNFSSVRLISSIRSLADCLPNVARSYLADL LPPCVRVVYLDSDLIIVDDIAKLAATDLGRDSVLAAPYCNANFTSYTSTFSWNP LSLTFADRKACFYNTGVMDLSLRWREGAYTSRIEENMAQKMRITVELGSLPPFLIV PAGLIKPVNHRWNGHGLGDNFRGLCRDLHPGPVSLHWSGKGKWARLDAGRPCPLD ALMAYPDLLPTPALDS"	
	complement(9620..10300)	
	/note="similar to Zn finger protein gi 6319911"	
	/codon_start=1	
	/evidence=not experimental	
	/product="F18014.3"	
	/protein_id="AAF79455.1"	
	/db_xref="GI:8778447"	
	/translation="MSDPVSCSGNDTNNDSNFECNLCLDIAQDPVTLGHGLFCW PCYKMLHLHSQSKCPCKVAYIEEDRLVLYGKSSADPRSCKIPGLEVNPSPSGO RPETAQPDNDHGFAGHHGFGMGCPAAMASARFQNTVLSAARGLLPSLNPRLHGH RPTQAPVNAAGAAASGCFPHGFSNPFHGSHHMSYQRHGRQGGQDHLHLLILLIVFVV VPSLFLUS"	
	join(12523..12577,12686..13374)	
	/note="hypothetical protein"	
	/codon_start=1	
	/evidence=not experimental	
	/product="F18014.4"	
	/protein_id="AAF79420.1"	
	/db_xref="GI:8778412"	
	/translation="MAIFSTSHLLFIFIIATCTISVSGTFTLTNHCSTIWPGLIT ANGAQLDGGGFALAGSSVFTVSPGMSGRFWARTYCNFDASGSGCTGDCGSKLKC AGAGGAPATLAEFTIGSGKNVADFYDVLVDGYNVQMKITPQGGSGDCKTAGCV SDVNAICPELQVTPSGVAACKSACEAFNKPEYCCCTGAYSTPATCPTNYSKIFPQCA CPASAYAYDDASSTFTCTNANVEISPCS"	
	complement(join(14020..14136,14236..14304,14402..14464, 14557..14696,14879..15005,16274..16486))	
	/note="unknown protein; similar to EST gb H76493.1"	
	/codon_start=1	
	/evidence=not experimental	
	/product="F18014.5"	
	/protein_id="AAF79454.1"	
	/db_xref="GI:8778446"	
	/translation="MLEAVDSSGVNNGGFPQIOSFYDCSSSEELSVLRHKVUVTG NNRKSVLVGLQGVKAVGLGVGHMVLVINGLEKQRNALSVLPFPFGNEEDDLDD FENTQRNGSDMTSEDTLTPHKSKRGQRSSRSHKTSKLSDSQSKSGFTIPPNM KVDLSKLEMPALLNWRHFLVDIAIPNSKEQIDIVQRHWSQQQDMLQVIVGFVQA AKRMKACKFSQKESRNTDLNCIS"	
	join(17994..18489,18568..18698,18814..18847,19006..19099, 19170..19347,19441..19479,19616..19715,19855..19994)	
	/note="hypothetical protein"	
	/codon_start=1	
	/evidence=not experimental	
	/product="F18014.6"	
	/protein_id="AAF79421.1"	
	/db_xref="GI:8778413"	
	/translation="MAKTDKLAQLDSGIYESDEFNWFLLDTVRIITNRSYTRFKVSPS AVYSRFFNSKQLNQHSGSESNPKRKRKQKNSSFLHPSVGEQASNLRHQEARLFLSKAH	

ESFLKTEILSLSLTKGLSDDDDDSSLLNKKCCDDVSEFIELGGVWQAPFVEITLTSFN  
 HCNESGSENRQVFOQPNNLVNEIEGEEVEAFENRNYIMPRNSCFYMSDLHINL  
 VPDNCEGYNLYVIDDPWENASQHSKFVILMYLPTLNQYFLSLPIKQLAHAEGL  
 VALVWTRKLLSFEKLEFPWGLIKYVATMYLKVKPDGTLICDLDLAGSEKSRDGF  
 KLADKNQIINSIPGDFSRKPPIGDILLKHTPPGSGPARCLELFAREMAAGWTSGNEPL  
 HFQDSRYFLKV"  
 join(22037, .22300, 22618, .23361)  
 /notes="unknown protein; similar to ESTs gb|A1992723.1,  
 gb|T41777.1, dbj|AV440980.1, gb|A1997341.1, and  
 dbj|J45440785.1"  
 /codon\_start=1  
 /evidence=not experimental  
 /product="F18014.7"  
 /protein\_id="AAF79422.1"  
 /db\_xref="GI:8778414"  
 /translations="MTSDGATSTASAAAAAAMATRRKPSWRRENNRRRRRAVAA  
 KIVTLAQGNVLPKHDNDNEVLKALCSAGWVVEDGTTYKKGHKPLPGDMAGSSS  
 RATPYSHNSQSLSSFPDSEILSYQVSPSSSPSPSRVGDPHNISTIPFLRNGGIP  
 SSUPLKRISNAPVTPVPSSTGRNPKPLTWESFTKQSPMAAKQSMSTLNTFPFYFA  
 SAPASPTHRQFHAPATIPCDSDSTVDSGHWISQKPAQQQPSFASGSMVPTSTFN  
 LVKPAPOOLSPNTAAIQEIQCSSEFKPENSQVKEGRIHDVAMEDLELTLLGNGKAH  
 S"  
 complement(join(23868, .24983, 25237, .25407))  
 /notes="unknown protein; similar to EST gb|F13858.1,  
 emb|Z38050.1, and emb|Z38087.1"  
 /codon\_start=1  
 /evidence=not experimental  
 /product="F18014.8"  
 /protein\_id="AAF79453.1"  
 /db\_xref="GI:8778445"  
 /translations="MAGRRDRSQOLRGSRITAIILGIFGCVCAVLFPVGFNNSSS  
 LKASEHLSKSNQVGSASPERVMLKSDFTVLSKNAELKKQRELTETKRLAEQ  
 GSDNARKVALGTQIKAGPFGTKSLRTPTLPDESINPRLAKILETAVDKVEIV  
 ALANARKVAMLEQIASIKRGTITNVLVALDDYIENLCKENDVAVYKRPDPKDVDT  
 GKTGGNHAVSGLAFRLVRLREFLOLGYGVLLSDVDIVFLQNPPHLYRDSVSMSDGH  
 NHTAYGNDVDFSPAGWARYAHTMRIWVFNSEFTALHASKRVMDMYEFMNSKVLFTKTRKHELKKVKVPI  
 WQDAVFNESELPFPMGYVETALHASKRVMDMYEFMNSKVLFTKTRKHELKKVKVPI  
 VHYNYHPDLNRMQAVVEFYVGGKDALDSFPDGS"  
 complement(26100, .27917)  
 /notes="unknown protein; similar to ESTs gb|T42717.1,  
 emb|F13886.1, and emb|F13885.1"  
 /codon\_start=1  
 /evidence=not experimental  
 /product="F18014.9"  
 /protein\_id="AAF79452.1"  
 /db\_xref="GI:8778444"  
 /translations="MADPLNGKSEFFICSLFSLFTLLFISPLYATSESPVIEDVSTDA  
 VSVSTREAVLHLNLEELVKNLUTELVANLDAKSLATPKERKEISVDDDIGEEKERG  
 RAKAFSTYKTSPTMSERFQTSVAKLNSDATINVLPRDFEGSSKYFAIGDSKGRVY  
 VFLRNGDLEIETFEVTPVAVMSVSVFKNSFVVTYHQNGAVALLHRIHEGNSGDS  
 WNSNSVMEHVGGKFDVDSADPVTLLVHHVGRVYILATDLSGLKTLVLTENTVYGS  
 VIPSRLVFLKQRLPLFTSSGAGSLDRSMKIRETECEGLNHSLARTYVFDAAERSK

Alignment Scores:  
 Pred. No.: 3,85e-73 Length: 120977  
 Score: 1137.50 Matches: 335  
 Percent Similarity: 40.58% Conservative: 68  
 Best Local Similarity: 33.74% Mismatches: 78  
 Query Match: 40.93% Indels: 518  
 DB: 8 Gaps: 18

US-10-051-909-36 (1-553) x AC025808 (1-120977)

Qy	52	Gluc1SerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAenThrGlySer	71
		:	
Db	67805	GAAGCGGAAGAATGAT-----CTTCGGCGACCGTCTTTACACACAGGAAGT	67755
Qy	72	TptTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
Db	67757	TGGTATCGGATGGGTTCCAGACATCATGATG-----TTGGNAATCGTCTCAAGTT	67700
Qy	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
		:	

67706	Db	ATTGCGAGACAGCTCAATTTCTGTCTTAGCTGTGGTTTTGATTGTTGGCTCTTGGTCTCTATT	6764
112	Qy	GlnPheGlyPheThrSer 	117
67646	Db	CAATTCCGATTCACTGT - AAGAACACATATAACAAATATCTCAACTTTCTTATATTTTT	67588
117	Qy	----- 	117
67587	Db	TTTATCAATGTGGATTCTCTGCTTTTGCTCAAAITTGACTAAGAAAGTGTGATTTTGAT	67528
118	Qy	-----GlyPheSerSerProThrGlnAspAlaMetValArgAsp 	130
67527	Db	TTATCTTTTGTTCAGTGTGGTTATCTCTCACCACTCAAGCTGCCATTACTAAGAT	67468
131	Qy	LeuAsnLeuSerIleSerGlu 	137
67467	Db	CTTTGGTTTAACGTATCAAGAGGTATGAAGGTCCCATCTTTATTAAAGAACTGTTCTGTTT	67408
137	Qy	----- 	137
67407	Db	TTTTTTTTTTTTTTGTAAGACTCTTGTTCTGTAAATTTTCTGACTTTTCTGTTCTTTGT	67348
138	Qy	-----PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla 	153
67347	Db	TTGATGTTTGTAGTACTCTGTGTTTGGGTCTCTATCCATGTGGGTGTATGTTTGAGCA	67288
154	Qy	IleAlaSerGlyGlnMetAlaGluIleGlyArg - LysGly 	167
67287	Db	ATTGCGAGTGGTCAGATTGCTGAATACGTTGGACGGAAAGGGGTATAATAGTTTGTGTTT	67228
167	Qy	----- 	167
67227	Db	TTGATTTCTAAACATGTTCTCATCTGTGCAAAATTTGTGATGTAACATCTTCTTCTTGT	67168
168	Qy	-----SerLeuMetIleAlaAlaIlePr 	175
67167	Db	GTATCTATTACTGTGTTTTCTTCTCTTTTGACAGCTCTCTGATGATGCTGCAATTC	67108
175	Qy	oAsnIleIleGlyTrioLeuAlaIleSerPheAla - Lys 	187
67107	Db	CAATATTATTGGATGCTTTTCGATATCATTTGCAAAAGTGAGTGTTTGGTGTCTCAAA	67048
187	Qy	----- 	187
67047	Db	GTTTCTGATATCTTGGCTTATCTCTATACCTACTAAAGTTGTTGTTCTTGGTTTTTG	66988
188	Qy	-----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyGlyPheGlyValGlyIleI 	206
66987	Db	TTTAGATATCTCTTTCTTTTACATGGGAAGATTGTTAGAAGCTTCGGAGTTGGGATTA	66928
206	Qy	IleSerTyrThrVal 	210
66927	Db	TCCTCTATACGGT - AATATTAGATTCACATATTTCTTGATTTTACGAATGTTGTTTGG	66869
211	Qy	-----ProValTyrIleAlaIag 	216
66868	Db	AGTGATCAATATAATTACACCTTTTGAATTGCGATTTTAGGTTCCCGTATATATAGCTG	66809
216	Qy	IuIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu 	232
66808	Db	AGATCGCTCCACAGACCATTGAGAGGAGCCTTAGGTTTCAGTTAAACAGGT - TTGT	66750
232	Qy	----- 	232
66749	Db	TAGTTAGCTTCACTATTTTATGTTTCAATTTTATGAGCTGATTCATATGTTAAACATAA	66690
233	Qy	-----SerValThrPheGlyIlePheLeuAlaTyrLeuL 	244
66689	Db	CAATAAATGTTTGTGTTTATAGCTTTCTGTAAACAATGGGATAATGCTGGCGTATTTAC	66630
244	Qy	eGlyMetPheIleProTrioArgLeuLeuAlaVal 	255
66629	Db	TCGGTCTTTGTTTCATGGAGAAATCTTCGATTTTGGGTAACTCACAAAACATTTTGA	66570





REFERENCE TITLE JOURNAL	4 (bases 1 to 68041) Town,C.D. and Kaul,S. Direct Submission Submitted (05-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
REFERENCE AUTHORS TITLE JOURNAL	5 (bases 1 to 68041) Town,C.D. and Kaul,S. Direct Submission Submitted (12-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
REFERENCE AUTHORS TITLE JOURNAL	6 (bases 1 to 68041) Town,C.D. PhD. Direct Submission Submitted (22-JAN-2001) On Jan 22, 2001 this sequence version replaced gi:12280794. Address all correspondence to:at@tigr.org
COMMENT	BAC clone P22H5 is from Arabidopsis thaliana chromosome 1 The orientation of the sequence is from SP6 to T7 end of the BAC clone. Genes were identified by a combination of several methods: Gene prediction programs including Genscan+ (Chris Burge, http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky, http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant of GlimmerM, see Mihaela Pertea, http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and GeneSplicer (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (http://www.tigr.org/cdb/tgi.shtml). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by repeatmasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html). Location/Qualifiers 1..68041 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="1" /clone="P22H5" complement(418..5404) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" /codon_start=1 /product="unknown protein, 3' partial; 5120-134" /protein_id="AAG12650.1" /db_xref="GI:1092277" /translation="MEYKSPPTATTSRKISNNRSHLSFSANAVDGVFSPVNSKS PLVDYGEIFRSGSPSSIPFLDPELVNKGKVDYRRSKLDYSSVFGGLGACDFAT PKEVITKSEKTIINEDKKRNRKGGNSDVLPCNEGKSPENVMRMKHSIDISYHQTVP RNEGATHLTQVATNPPIPTQVTDNTSLHKIESKSTPIPAVEKKLPCNEGEVVK ASRQSKTEVDENI FARDGCTSDRDTCTKTVNGEYRDVKKPSSFOCTLNGEHGAS ERLGLNSGSERYETEDADSPSPYFDAETDENSVAESSAALKKAIEAQIRNMI AQGMERKKSGRFCAKLSGCDSDSKENKTKVEGTIESRNNNSQILGEMVKPSEQ SPFNEGDQKARKLWGPGLKSTSDHKPELEBEDIITLIEEQARRGRKHWP GGIFKSVMSKQEPENLPAKPEPTTKQEVQITENPFYFQLGSLKCVVEAFTG SKVSKQDEKQFTKENSITVQWDESDSQEMLAGIPVLETYLRVEBTPQOETESK EMNIEKSESTICAFTERSQNMKEGTGWKQVSKAFCKEDGSGVKDQENGDPHCNVD QGEKEIVSEQMLVGPDDSKTVVREVESTPTPSLNNKTRPQWLESAMVSFNRVNSE PGNIDVQEAHVHVRPRRRVWKTSEDVYNNIKAPKGNRNPQWLESAMVETAMSFEH
FEATURES source	1..68041 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /cultivar="Columbia" /db_xref="taxon:3702" /chromosome="1" /clone="P22H5" complement(418..5404) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" /codon_start=1 /product="unknown protein, 3' partial; 5120-134" /protein_id="AAG12650.1" /db_xref="GI:1092277" /translation="MEYKSPPTATTSRKISNNRSHLSFSANAVDGVFSPVNSKS PLVDYGEIFRSGSPSSIPFLDPELVNKGKVDYRRSKLDYSSVFGGLGACDFAT PKEVITKSEKTIINEDKKRNRKGGNSDVLPCNEGKSPENVMRMKHSIDISYHQTVP RNEGATHLTQVATNPPIPTQVTDNTSLHKIESKSTPIPAVEKKLPCNEGEVVK ASRQSKTEVDENI FARDGCTSDRDTCTKTVNGEYRDVKKPSSFOCTLNGEHGAS ERLGLNSGSERYETEDADSPSPYFDAETDENSVAESSAALKKAIEAQIRNMI AQGMERKKSGRFCAKLSGCDSDSKENKTKVEGTIESRNNNSQILGEMVKPSEQ SPFNEGDQKARKLWGPGLKSTSDHKPELEBEDIITLIEEQARRGRKHWP GGIFKSVMSKQEPENLPAKPEPTTKQEVQITENPFYFQLGSLKCVVEAFTG SKVSKQDEKQFTKENSITVQWDESDSQEMLAGIPVLETYLRVEBTPQOETESK EMNIEKSESTICAFTERSQNMKEGTGWKQVSKAFCKEDGSGVKDQENGDPHCNVD QGEKEIVSEQMLVGPDDSKTVVREVESTPTPSLNNKTRPQWLESAMVSFNRVNSE PGNIDVQEAHVHVRPRRRVWKTSEDVYNNIKAPKGNRNPQWLESAMVETAMSFEH
gene	complement(418..5404) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" /codon_start=1 /product="unknown protein, 3' partial; 5120-134" /protein_id="AAG12650.1" /db_xref="GI:1092277" /translation="MEYKSPPTATTSRKISNNRSHLSFSANAVDGVFSPVNSKS PLVDYGEIFRSGSPSSIPFLDPELVNKGKVDYRRSKLDYSSVFGGLGACDFAT PKEVITKSEKTIINEDKKRNRKGGNSDVLPCNEGKSPENVMRMKHSIDISYHQTVP RNEGATHLTQVATNPPIPTQVTDNTSLHKIESKSTPIPAVEKKLPCNEGEVVK ASRQSKTEVDENI FARDGCTSDRDTCTKTVNGEYRDVKKPSSFOCTLNGEHGAS ERLGLNSGSERYETEDADSPSPYFDAETDENSVAESSAALKKAIEAQIRNMI AQGMERKKSGRFCAKLSGCDSDSKENKTKVEGTIESRNNNSQILGEMVKPSEQ SPFNEGDQKARKLWGPGLKSTSDHKPELEBEDIITLIEEQARRGRKHWP GGIFKSVMSKQEPENLPAKPEPTTKQEVQITENPFYFQLGSLKCVVEAFTG SKVSKQDEKQFTKENSITVQWDESDSQEMLAGIPVLETYLRVEBTPQOETESK EMNIEKSESTICAFTERSQNMKEGTGWKQVSKAFCKEDGSGVKDQENGDPHCNVD QGEKEIVSEQMLVGPDDSKTVVREVESTPTPSLNNKTRPQWLESAMVSFNRVNSE PGNIDVQEAHVHVRPRRRVWKTSEDVYNNIKAPKGNRNPQWLESAMVETAMSFEH
mRNA	complement(418..5404) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" /codon_start=1 /product="unknown protein, 3' partial; 5120-134" /protein_id="AAG12650.1" /db_xref="GI:1092277" /translation="MEYKSPPTATTSRKISNNRSHLSFSANAVDGVFSPVNSKS PLVDYGEIFRSGSPSSIPFLDPELVNKGKVDYRRSKLDYSSVFGGLGACDFAT PKEVITKSEKTIINEDKKRNRKGGNSDVLPCNEGKSPENVMRMKHSIDISYHQTVP RNEGATHLTQVATNPPIPTQVTDNTSLHKIESKSTPIPAVEKKLPCNEGEVVK ASRQSKTEVDENI FARDGCTSDRDTCTKTVNGEYRDVKKPSSFOCTLNGEHGAS ERLGLNSGSERYETEDADSPSPYFDAETDENSVAESSAALKKAIEAQIRNMI AQGMERKKSGRFCAKLSGCDSDSKENKTKVEGTIESRNNNSQILGEMVKPSEQ SPFNEGDQKARKLWGPGLKSTSDHKPELEBEDIITLIEEQARRGRKHWP GGIFKSVMSKQEPENLPAKPEPTTKQEVQITENPFYFQLGSLKCVVEAFTG SKVSKQDEKQFTKENSITVQWDESDSQEMLAGIPVLETYLRVEBTPQOETESK EMNIEKSESTICAFTERSQNMKEGTGWKQVSKAFCKEDGSGVKDQENGDPHCNVD QGEKEIVSEQMLVGPDDSKTVVREVESTPTPSLNNKTRPQWLESAMVSFNRVNSE PGNIDVQEAHVHVRPRRRVWKTSEDVYNNIKAPKGNRNPQWLESAMVETAMSFEH
CDS	complement(418..5404) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" complement(join(<418..594,684..773,873..938,1034..1095, 1179..1227,1319..4820,5040..5404)) /gene="F22H5.15" /codon_start=1 /product="unknown protein, 3' partial; 5120-134" /protein_id="AAG12650.1" /db_xref="GI:1092277" /translation="MEYKSPPTATTSRKISNNRSHLSFSANAVDGVFSPVNSKS PLVDYGEIFRSGSPSSIPFLDPELVNKGKVDYRRSKLDYSSVFGGLGACDFAT PKEVITKSEKTIINEDKKRNRKGGNSDVLPCNEGKSPENVMRMKHSIDISYHQTVP RNEGATHLTQVATNPPIPTQVTDNTSLHKIESKSTPIPAVEKKLPCNEGEVVK ASRQSKTEVDENI FARDGCTSDRDTCTKTVNGEYRDVKKPSSFOCTLNGEHGAS ERLGLNSGSERYETEDADSPSPYFDAETDENSVAESSAALKKAIEAQIRNMI AQGMERKKSGRFCAKLSGCDSDSKENKTKVEGTIESRNNNSQILGEMVKPSEQ SPFNEGDQKARKLWGPGLKSTSDHKPELEBEDIITLIEEQARRGRKHWP GGIFKSVMSKQEPENLPAKPEPTTKQEVQITENPFYFQLGSLKCVVEAFTG SKVSKQDEKQFTKENSITVQWDESDSQEMLAGIPVLETYLRVEBTPQOETESK EMNIEKSESTICAFTERSQNMKEGTGWKQVSKAFCKEDGSGVKDQENGDPHCNVD QGEKEIVSEQMLVGPDDSKTVVREVESTPTPSLNNKTRPQWLESAMVSFNRVNSE PGNIDVQEAHVHVRPRRRVWKTSEDVYNNIKAPKGNRNPQWLESAMVETAMSFEH
gene	complement(9354..10899) /gene="F22H5.16" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<9354..9638,9750..9953,10043..10328, 10461..10595,10671..>10899)) /gene="F22H5.16" complement(join(9481..9638,9750..9953,10043..10328, 10461..10595,10671..10856)) /gene="F22H5.16" /codon_start=1 /product="NADPH oxidoreductase, putative; 10572-9197" /protein_id="AAG12695.1" /db_xref="GI:10092282" /translation="MTTEKSKILVIGTGYMGFEFVIGSAKAGNPTFFALVREASLSDP /franktione="MTTEKSKILVIGTGYMGFEFVIGSAKAGNPTFFALVREASLSDP VSKITQSPFDGLVITLHGDNDHESLVAIKQVDVIVTIGHKQIFDQTKIISAIKE AGNKRFLPAEFGIDVTSVAPPAKSLFAGVQIRRAIEAGIPIYTVNVCAGGY LTLQFESGLISHTRDKAIFGDKNVPKDVITLGDGNKAVIYNKEEDVAAVMYKA VDLRLTKLTISSPNNILSNMWTLMKEKIGKSLKTHISEEQILKSIQVPIVDF KSNINAVFVGQDTSTPIEPWFGEASVLPDKVYTSIDELYSQFT" complement(11122..12573) /gene="F22H5.18" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<11122..11422,11516..11734,11791..12034, 12117..12251,12333..>12573)) /gene="F22H5.18" complement(join(11235..11422,11516..11734,11791..12034, 12117..12251,12333..12518)) /gene="F22H5.18" /codon_start=1 /product="NADPH oxidoreductase, putative; 12234-10951" /protein_id="AAG12680.1" /db_xref="GI:10092267" /translation="MASEKSKILVIGTGHIGKLIIEASVKAGHSTLALVREASLSDP NKGKTVQNEKDFQVTLHGDNDHESLVAIKQVDVIVTIGSMQILDQTKIISAIKE AGNKRFLPSEFGMDVKSSAVEPAKSAFGRKLQTRDIEAGIPIYTVNVCAGGY LPTLVQLEPLTSPPRDKVIFGDKNVRVEYFIIVINKBEDIAAYTIKAVDDPRTLNK TLIYNPNNTLSMNEIVTLWEKIGKSLKTHISEEQIFKSIQESVPPVNVLLSINHA VFVKGDTNFTIPEPSFGFEASLYPDIKYTSIDELYSFALGTSNT" complement(12872..14426) /gene="F22H5.17" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<12872..13216,13298..13513,13578..13821, 13974..14108,14193..>14426)) /gene="F22H5.17" complement(join(13053..13216,13298..13513,13578..13821, 13974..14108,14193..14378)) /gene="F22H5.17" /codon_start=1 /product="NADPH oxidoreductase, putative; 14094-12769" /protein_id="AAG12677.1" /db_xref="GI:10092264" /translation="MATEKSKILVIGTGYIGKFLVESA KAGHSTFALVREATLSDP VKGKTVQSPDFGLVITLHGDNDHESLVAIKQVDVIVTIGSMQILDQTKIISAIKE AGNKRFLPSEFGMDVKSSAVEPAKSAFAGIKIRRTIEAGIPIYTVNVCAGGY LPTLVQLEPLTSPPRDKVITLGDGNKVEYFAVINKBEDIAAYTIKAVDDPRTLNK LTIKSNNTLSMNEIVTLWEKIGKSLKTHISEEQILKSIQESVPPVNVLLSINHA FVNGDNTNISIEPSFGFEASLYPDVYKTSIDELYSFA"
mRNA	complement(9354..10899) /gene="F22H5.16" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<9354..9638,9750..9953,10043..10328, 10461..10595,10671..>10899)) /gene="F22H5.16" complement(join(9481..9638,9750..9953,10043..10328, 10461..10595,10671..10856)) /gene="F22H5.16" /codon_start=1 /product="NADPH oxidoreductase, putative; 10572-9197" /protein_id="AAG12695.1" /db_xref="GI:10092282" /translation="MTTEKSKILVIGTGYMGFEFVIGSAKAGNPTFFALVREASLSDP /franktione="MTTEKSKILVIGTGYMGFEFVIGSAKAGNPTFFALVREASLSDP VSKITQSPFDGLVITLHGDNDHESLVAIKQVDVIVTIGHKQIFDQTKIISAIKE AGNKRFLPAEFGIDVTSVAPPAKSLFAGVQIRRAIEAGIPIYTVNVCAGGY LTLQFESGLISHTRDKAIFGDKNVPKDVITLGDGNKAVIYNKEEDVAAVMYKA VDLRLTKLTISSPNNILSNMWTLMKEKIGKSLKTHISEEQILKSIQVPIVDF KSNINAVFVGQDTSTPIEPWFGEASVLPDKVYTSIDELYSQFT" complement(11122..12573) /gene="F22H5.18" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<11122..11422,11516..11734,11791..12034, 12117..12251,12333..>12573)) /gene="F22H5.18" complement(join(11235..11422,11516..11734,11791..12034, 12117..12251,12333..12518)) /gene="F22H5.18" /codon_start=1 /product="NADPH oxidoreductase, putative; 12234-10951" /protein_id="AAG12680.1" /db_xref="GI:10092267" /translation="MASEKSKILVIGTGHIGKLIIEASVKAGHSTLALVREASLSDP NKGKTVQNEKDFQVTLHGDNDHESLVAIKQVDVIVTIGSMQILDQTKIISAIKE AGNKRFLPSEFGMDVKSSAVEPAKSAFGRKLQTRDIEAGIPIYTVNVCAGGY LPTLVQLEPLTSPPRDKVIFGDKNVRVEYFIIVINKBEDIAAYTIKAVDDPRTLNK TLIYNPNNTLSMNEIVTLWEKIGKSLKTHISEEQIFKSIQESVPPVNVLLSINHA VFVKGDTNFTIPEPSFGFEASLYPDIKYTSIDELYSFALGTSNT" complement(12872..14426) /gene="F22H5.17" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<12872..13216,13298..13513,13578..13821, 13974..14108,14193..>14426)) /gene="F22H5.17" complement(join(13053..13216,13298..13513,13578..13821, 13974..14108,14193..14378)) /gene="F22H5.17" /codon_start=1 /product="NADPH oxidoreductase, putative; 14094-12769" /protein_id="AAG12677.1" /db_xref="GI:10092264" /translation="MATEKSKILVIGTGYIGKFLVESA KAGHSTFALVREATLSDP VKGKTVQSPDFGLVITLHGDNDHESLVAIKQVDVIVTIGSMQILDQTKIISAIKE AGNKRFLPSEFGMDVKSSAVEPAKSAFAGIKIRRTIEAGIPIYTVNVCAGGY LPTLVQLEPLTSPPRDKVITLGDGNKVEYFAVINKBEDIAAYTIKAVDDPRTLNK LTIKSNNTLSMNEIVTLWEKIGKSLKTHISEEQILKSIQESVPPVNVLLSINHA FVNGDNTNISIEPSFGFEASLYPDVYKTSIDELYSFA"
CDS	complement(9354..10899) /gene="F22H5.16" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<9354..9638,9750..9953,10043..10328, 10461..10595,10671..>10899)) /gene="F22H5.16" complement(join(9481..9638,9750..9953,10043..10328, 10461..10595,10671..10856)) /gene="F22H5.16" /codon_start=1 /product="NADPH oxidoreductase, putative; 10572-9197" /protein_id="AAG12695.1" /db_xref="GI:10092282" /translation="MTTEKSKILVIGTGYMGFEFVIGSAKAGNPTFFALVREASLSDP /franktione="MTTEKSKILVIGTGYMGFEFVIGSAKAGNPTFFALVREASLSDP VSKITQSPFDGLVITLHGDNDHESLVAIKQVDVIVTIGHKQIFDQTKIISAIKE AGNKRFLPAEFGIDVTSVAPPAKSLFAGVQIRRAIEAGIPIYTVNVCAGGY LTLQFESGLISHTRDKAIFGDKNVPKDVITLGDGNKAVIYNKEEDVAAVMYKA VDLRLTKLTISSPNNILSNMWTLMKEKIGKSLKTHISEEQILKSIQVPIVDF KSNINAVFVGQDTSTPIEPWFGEASVLPDKVYTSIDELYSQFT" complement(11122..12573) /gene="F22H5.18" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<11122..11422,11516..11734,11791..12034, 12117..12251,12333..>12573)) /gene="F22H5.18" complement(join(11235..11422,11516..11734,11791..12034, 12117..12251,12333..12518)) /gene="F22H5.18" /codon_start=1 /product="NADPH oxidoreductase, putative; 12234-10951" /protein_id="AAG12680.1" /db_xref="GI:10092267" /translation="MASEKSKILVIGTGHIGKLIIEASVKAGHSTLALVREASLSDP NKGKTVQNEKDFQVTLHGDNDHESLVAIKQVDVIVTIGSMQILDQTKIISAIKE AGNKRFLPSEFGMDVKSSAVEPAKSAFGRKLQTRDIEAGIPIYTVNVCAGGY LPTLVQLEPLTSPPRDKVIFGDKNVRVEYFIIVINKBEDIAAYTIKAVDDPRTLNK TLIYNPNNTLSMNEIVTLWEKIGKSLKTHISEEQIFKSIQESVPPVNVLLSINHA VFVKGDTNFTIPEPSFGFEASLYPDIKYTSIDELYSFALGTSNT" complement(12872..14426) /gene="F22H5.17" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<12872..13216,13298..13513,13578..13821, 13974..14108,14193..>14426)) /gene="F22H5.17" complement(join(13053..13216,13298..13513,13578..13821, 13974..14108,14193..14378)) /gene="F22H5.17" /codon_start=1 /product="NADPH oxidoreductase, putative; 14094-12769" /protein_id="AAG12677.1" /db_xref="GI:10092264" /translation="MATEKSKILVIGTGYIGKFLVESA KAGHSTFALVREATLSDP VKGKTVQSPDFGLVITLHGDNDHESLVAIKQVDVIVTIGSMQILDQTKIISAIKE AGNKRFLPSEFGMDVKSSAVEPAKSAFAGIKIRRTIEAGIPIYTVNVCAGGY LPTLVQLEPLTSPPRDKVITLGDGNKVEYFAVINKBEDIAAYTIKAVDDPRTLNK LTIKSNNTLSMNEIVTLWEKIGKSLKTHISEEQILKSIQESVPPVNVLLSINHA FVNGDNTNISIEPSFGFEASLYPDVYKTSIDELYSFA"

gene 15081..16153  
/gene="P22H5.1"  
/note="similar to GI:6939839 from [Oryza sativa]"  
mRNA join(15081..15228,15380..15470,15661..16153)  
/gene="P22H5.1"  
CDS join(15171..15228,15380..15470,15661..16153)  
/gene="P22H5.1"  
/codon\_start=1  
/product="GSH-dependent dehydroascorbate reductase 1,  
putative; 14887-15869"  
protein\_id="AAG12679.1"  
db\_xref="GI:10092266"  
/translation="MALDIQVAVAGDVLGDCPPSORVLLTLEKKLKYKTHLNV  
SNPQWFLIDISPEQKVPVQLDGWVADSVIQLLEKYPESLKTTPPEFASVSGSKI  
FGAFVFLKSDANDSGSKALVDELEALENLKTHSGFPVAGKIFATVDLSLAPKLYH  
LEVALGHTKNSVPESLTSVRNYAKALFRESFENTKAKEIIVVAGHESKUNA"  
complement(16410..18554)  
gene /gene="P22H5.2"  
complement(join(16410..16558,16703..16830,17050..18554))  
mRNA /gene="P22H5.2"  
complement(join(16410..16558,16703..16830,17050..18554))  
CDS /gene="P22H5.2"  
/codon\_start=1  
/product="unknown protein; 18270-16126"  
protein\_id="AAG12682.1"  
db\_xref="GI:10092269"  
/translation="MPPFRFRIPFSSNSPSRISSTSSPSPPPTPPSRPPFRPAGIA"  
Alignment Scores:  
Pred. No.: 5,46e-72 Length: 68041  
Score: 1117.50 Matches: 340  
Percent Similarity: 38.10% Conservative: 73  
Best Local Similarity: 34.37% Mismatches: 82  
Query Match: 40.21% Indels: 595  
DB: 8 Gaps: 18  
US-10-051-909-36 (1-553) x AC025814 (1-68041)  
QY 52 GluSerGlySerAspHisValLeuArgArgProLeuLeuAsnThrGlySer 71  
Db 33823 GAGGAGGCTAAGATGAT-----TTGGCTGCGCGCTTTATTACACCGGAAGC 33870  
QY 72 TrpTyrMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91  
Db 33871 TGGTACCGGATGGGTTCGAGACAATCGAGTATG-----ATGGGTTCCTCTCAAGTT 33921  
QY 92 LeuAtGluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProile 111  
Db 33922 ATTGAGAGAGCTATCTCTGTTCTTGATGCTGTTTGTGATGCTGTTGCTTGGCCCTATT 33981  
QY 112 GlnPheGlyPheThr----- 116  
Db 33982 CAATTCGGATTCACTGTAGACAAATTCAAATCTCAACTTTTATTATTATTTATTTCC 34041  
QY 116 ----- 116  
Db 34042 TTTGTAGAGAAAGGTGTGACTTTTGGTGGAAAACTAGTGTGAGCTGGATTGATGATC 34101  
QY 116 ----- 116  
Db 34102 TTGGGTTTGGTTTGTATTGTCATTTCATGCAAGTCTGGTGATTGGAATCTCATATT 34161  
QY 116 ----- 116  
Db 34162 TGGTGAACGAGATGATTTTTTCAAGTTTGTCTGACAGATTAGAGGACTTTGAAATTGGGA 34221  
QY 116 ----- 116  
Db 34222 ATTTCTGATTGTTTGGCTTAAACCTTCATGTTTGTGATTTTGATCACCATCTTTAGTACTC 34281  
QY 117 -----  
Db 34282 TGTGTTTGGTCTTAAACCTTCATCCCTTTTGTTCAGTGTGGTTATTTCATCTCCAACTCA 34341

QY 124 nAspAlaMetValArgArgAspLeuAsnLeuSerSerLeuSerGlu----- 137  
Db 34342 AGCTGCAATCACCAGGATCTCGGTTTAAACAGTATCTGAGGTAATGAGCTGATCAATTGTC 34401  
QY 137 ----- 137  
Db 34402 TAGAGCTGCTGGTTTTTATCATTTGTTTGTAGCCGTTTGTGATTTGTGATGTACCACTC 34461  
QY 138 -----PheSerAlaPheGlySerLeuSerAsnValGlyMetValGlyAlaI 154  
Db 34462 CATGTTTTAGTACTCTGTGTTTGGTTCTCTTCCAAATGTGGTCTGATGTTGTTGCAAT 34521  
QY 154 eAlaSerGlyGlnMetAlaGluTyrIleGlyArg-LysGly----- 167  
Db 34522 TGCCATGTGTGAGATTGAGAAATACATTGGACGAAAGGGGTATACTTTGTATTCTGATT 34581  
QY 167 ----- 167  
Db 34582 CTTGATCGTTCTGTTTTAGATCCATTTTATAAAGTCCTCTTCTTGGCTTCAACATCT 34641  
QY 168 -----SerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpL 181  
Db 34642 TATTGTGTTTTCTGGACAGTCTCTTATGATGTCTGCAATTCCTCAATATCAATGATGCG 34701  
QY 181 euAlaIleSerPheAlaLys----- 187  
Db 34702 TTTGCATATCATTTTGC-AAAAAGTCAGTTTTTGTCTCTCATACCAATAAATTTGCGGATT 34760  
QY 188 -----AspAlaSerPheL 192  
Db 34761 GATTGTGTTGCTGATTTAACTCTTTTACATCAATCTGTGTTTGTAGGATACTCTTTTTC 34820  
QY 192 euTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr----- 209  
Db 34821 TTTACATGGGAAGACTTTTGAAGGCTTTGGCGTTGGGATCATCTTTACACTGTAATAC 34880  
QY 209 ----- 209  
Db 34881 TGAACTTAACTAGCTATTATTACTTATGGAAATACGGTTTCAGTATATAATGCAACTCGGTT 34940  
QY 210 -----ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaL 226  
Db 34941 CGTGATTTTGTGATCTGTATATATCGCTGAGATAGCTCCAAATAATATGAGAGGAT 35000  
QY 226 euGlySerValAsnGlnLeu----- 232  
Db 35001 TGGGTTCGGTTAACCAAGGT-TAGACATGTGTAGTGAATAATTCCTGTTTGTGCTTTGCTG 35059  
QY 233 -----SerValT 235  
Db 35060 TACTGACTCGTAATGCTAAACAAGAAATGTTGTGTAACCTTCTGCTGTAGCTTTCTGTGA 35119  
QY 235 hrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuAlaV 255  
Db 35120 CAATTGGAATAATGTTGGCGTATTATTCTTGGCTCTTTTGTTCATGGAGAATTTCTGCAG 35179  
QY 255 al----- 255  
Db 35180 TTCTGGTAAATCTCATTTACAGCGCTGAGACTATTTTGTGATTATGAATAATTGCTTAC 35239  
QY 256 -----IleGlyAlaLeuProCysThrMetL 264  
Db 35240 GCTTAATCACTTCTGTTCTCTTCTTCTTATCTTATAGGATATGCGGTGTACATTAT 35299  
QY 264 euIleProGlyLeuPhePheIleProGluSerProArgTyrIle----- 278  
Db 35300 TGATACCAAGGTCTCTTTTTCATCCCGGAATCCCTCGCTGGTGGTAAAGCCCTTACAGTAG 35359  
QY 278 ----- 278  
Db 35360 TTTGCTTATTCTTCTTGTAGTTAGTGATATTTTGGTTACTCCACTCTTTTTTTTTTTTTT 35419

```
QY 279 -----AlaLeuMet 281
Db 35420 TAATAACATCAATTAATCATCTCTGATCGGATACATTGATTCGTTAAATGTAGGCAAAA 35479
QY 281 etAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspI 301
Db 35480 TGGGTATGACAGATGAATTTGAACTTCATTACAAAGTTCTTCGAGGATTTGAGACTGATA 35539
QY 301 leThrThrGluValAsnAspIleLeu 309
Db 35540 TTACCGTTGAAGTTAATGAATAATCAAGGTAACATACCACTTAAATCTTTCCTTCCAAGA 35599
QY 310 -----Arg 310
Db 35600 GTTACATGAAGACTGGTTTCAACAGCTAATTTTCATCTTTCGCAATCTCGCAGAGA 35659
QY 311 AlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLys 330
Db 35660 TCTGTGGCATCACTCAAGAGCGAATACAGTTCGGTTTGTAGATCTCAACACGAGGAGA 35719
QY 331 TyrArgThrProLeuLeu 336
Db 35720 TACTATTCCCACTTATGGTACAAATATTGAATATCTTCTCGAATATACCTTTGTCTA 35779
QY 337 -----LeuG1 338
Db 35780 TCATAAGAAACATTATCTTAAAACTCTCTTTTTCGTGTGTGTGTGTGTAAACAGGTGG 35839
QY 338 YileGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSe 358
Db 35840 TATAGGGTGTCTGTACTTCAACACTTGTGTGNAATTAATGGTGTCTTATCTATTCCAG 35899
QY 358 rSerIlePheLysAlaala 364
Db 35900 TACAATATTGAATCTGCAGGTAGGTTTCTTAACACCTGAAGCTTCTGTGTTTCATATT 35959
QY 365 -----GlyValThrAsnSerAspLeuA 372
Db 35960 ATTACTTTAATTTTTCATTTTCGTTGTTTCATGTTTAGGAGTTTACATCGAGTAACGAC 36019
QY 372 laThrCysSerLeuGlyValaileGlnValLeuAlaThrGlyValThrThrTrieLeuLeuA 392
Db 36020 CAACATTTGGGGTTCGGGGCTATTTCAGGTAGTACGACTGCATATCCACTTTGGTTGGTG 36079
QY 392 spArgAlaGlyArgArgIleLeuLeuIle 401
Db 36080 ACAAGCAGGTCTGCGGCTTCTGCTTACTGTGAGTATTAACTCATTTTTCATATCTTT 36139
QY 401 ----- 401
Db 36140 CCGACTTCCCTCTACAGAAAGATTTAAGACATATCCAAGATTAACATAATGCAAAATGGAAA 36199
QY 402 -----IleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValP 418
Db 36200 ATATTTCCAGATCTCTTCGGTTGGGATGACGATTAGCCTTGTAATGTTGTCAGCTGCTT 36259
QY 418 hePheLeuLys 421
Db 36260 TCTATCTTAAGGTACTTTTAAATCTTGAACCTTCCATACACATCATTTTCATCTTCTGCA 36319
QY 422 -----AspAsnIleSerG 426
Db 36320 ACAAAATCCAAGATTCGACTTTCTGAGTTTTCATCTCTCTTTTAGGAATTT-GTGTCTC 36378
QY 426 lnAspSerAsnSerTyrTrieLeuThrMetIleSerLeuValGlyIleVal 443
Db 36379 CTGATTCAGACATGTACAGTTGGCTGAGCATATTGTGAGTAGTGTGGAGTTGTGGTAAGTT 36438
QY 443 ----- 443
Db 36439 AGTTATATAGCGNACTTTCTAGAAATGCAATGAATGATTATTGTGCCTTAACTCTGTGG 36498
QY 444 -----SerPheValIleThrPheSerPheGlyMetGlyAlaIleProThrLeuM 460
```

```
Db 36499 AATCGTTATAGGCAATGGTTGTTCTTTTCTCATTTGGGAATGGGACCAATACCGTGGCTCA 36558
QY 460 etMetSerGluIle 464
Db 36559 TTATGCTCGAGGT-AATGAGTTGGTTAGAAATTTAGAACTAAGTGTGTGTAAGAACCGAAAA 36617
QY 465 -----L 465
Db 36618 GCGAAAAACTTATTGAACCTCTACCTTTGCTTGTCTTTGTTGTTGATATAAACAGATCC 36677
QY 465 euProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrieLeuThrs 485
Db 36678 TTCTGTGAACATPAAGGGTTTAGCTGGAAGTATTGCAACTCTAGCCAAITGGTTCTTTT 36737
QY 485 erPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrieSerValGlyGlyThrPheL 505
Db 36738 CTTGGTTGATCACCATGACAGCAAAATTTGCTGTAGCTGGACGAGTGGAGGTTCTCTTC 36797
QY 505 euSerTyrMet 508
Db 36798 TCTCTTTCTC-TTCTTTTCTCTTCTCTCATACCATTTGGGACAATAGAGCATTTATCTTG 36856
QY 509 -----ValValSerAlap 513
Db 36857 ATCGATTTTGGGTTAATATTGGACAGGAACCTTTCATCTCTGTATGAGTTGTTGTGCAT 36916
QY 513 heThrIleValPheValValLeuTrieValProGluThrLysGly***AsnSerArgGlyA 533
Db 36917 TCACAGTGGTGTTCGTGACTCTATGGTTCCTGAGACCAAGGCAA-AACTCTTTGAAGAA 36975
QY 533 spThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTrieLeuProGluCysLeuS 553
Db 36976 CTTCAATCCTTGTTCAGATGAACAAATTTGAACAACTTCAATTTTGTGTACCCCTCTCT 37035
QY 553 er 553
Db 37036 CC 37037
```

Search completed: April 13, 2005, 22:00:41

Job time : 5725.66 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 12:06:34 ; Search time 639.423 Seconds  
(without alignments)  
5127.663 Million cell updates/sec

Title: US-10-051-909-36  
Perfect score: 2779  
Sequence: 1 PSSSSFRPAGKKKKKQ.....TIFVSLSIQRQLWLPECLLS 553

Scoring table:  
BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10051909/runat\_13042005\_074037\_14040/app.query.fasta\_1.1678  
-DB=N\_Geneseq\_16Dec04 -QPM=fastap -SUFFIX=rng -MINMATCH=0\_1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USPR=US10051909 @CGN 1 1 891 @runat\_13042005\_074037\_14040 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	2777	99.9	Adg47939 Corn Beta
2	1710	61.5	Ada69666 Rice gene
3	1577	56.7	Ada69667 Rice gene
4	1182.5	42.6	Adr59784 Cotton cd
5	951.5	34.2	AAC36319 Arabidops

6	946.5	34.1	990	13	ADR59785	Adr59785 Cotton cd
7	943.5	34.0	1683	3	AAC39099	Aac39099 Arabidops
8	933	33.6	1395	3	AAC45857	Aac45857 Arabidops
9	923	33.2	1685	3	AAC45298	Aac45298 Arabidops
10	923	33.2	1743	3	AAC45290	Aac45290 Arabidops
11	920.5	33.1	1473	6	ABZ14526	Abz14526 Arabidops
12	916	33.0	1755	3	AAC40459	Aac40459 Arabidops
13	899	32.3	1359	6	ABZ13822	Abz13822 Arabidops
14	893	32.1	1768	12	ADM47939	Adm47939 Polynucle
15	877.5	31.6	1407	6	ABZ13486	Abz13486 Arabidops
16	877.5	31.6	1407	8	ADA67841	Ada67841 Arabidops
17	877.5	31.6	1642	3	AAC46009	Aac46009 Arabidops
18	868.5	31.3	1425	3	AAC50050	Aac50050 Arabidops
19	853.5	30.7	1425	3	AAC50049	Aac50049 Arabidops
20	837.5	30.1	1209	8	ADA70113	Ada70113 Rice gene
21	824	29.7	1452	3	AAC42659	Aac42659 Arabidops
22	805.5	29.0	1695	3	AAC49517	Aac49517 Arabidops
23	790.5	28.4	1436	3	AAC45400	Aac45400 Arabidops
24	786.5	28.3	1440	3	AAC33666	Aac33666 Arabidops
25	777	28.0	1374	6	ABZ13215	Abz13215 Arabidops
26	758.5	27.3	1378	3	AAC35123	Aac35123 Arabidops
27	754.5	27.2	1344	8	ADA69777	Ada69777 Rice gene
28	688.5	24.8	1724	4	ABL13389	Ab113389 Drosophil
29	682.5	24.6	1239	3	AAC34742	Aac34742 Arabidops
30	674.5	24.3	2072	4	AAF55867	Aaf55867 Murine GL
31	674	24.3	2217	4	AAF55865	Aaf55865 Human GLU
32	672	24.2	898	10	ADD17645	Add17645 DNA (seq
33	672	24.2	2087	4	AAF55866	Aaf55866 Rat GLUTX
34	669.5	24.1	2080	4	AAD09552	Aad09552 Human tra
35	663	23.9	1473	6	ABZ13864	Abz13864 Arabidops
36	663	23.9	1473	6	ADG88039	Adg88039 A. thalia
37	658.5	23.7	1445	8	ABZ24794	Abz24794 Human sol
38	658.5	23.7	1856	8	ABZ24792	Abz24792 Human sol
39	657.5	23.7	600	13	ACN59223	Acn59223 Cotton gy
40	652.5	23.5	1873	8	ABZ24793	Abz24793 Human sol
41	652.5	23.5	2514	4	ABL17067	Ab117067 Drosophil
42	640	23.0	1412	4	ABL13883	Ab113883 Drosophil
43	623.5	22.4	2487	12	ABZ24790	Abz24790 Human sol
44	623.5	22.4	2487	12	ADQ86239	Adq86239 Human tum
45	623.5	22.4	2487	13	ADR25824	Adr25824 Breast ca

ALIGNMENTS

RESULT 1  
ADG47939  
ID ADG47939 standard; cdna; 1953 BP.  
XX  
AC ADG47939;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Corn Beta-vulgaris-like sugar transport protein cDNA #4.  
XX  
KW Arabidopsis-like sugar transport protein;  
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;  
KW corn; plant; gene; ss.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1562  
FT /\*tag= a  
FT /product= "Corn Beta-vulgaris-like sugar transport  
FT protein"  
FT /note= "Xaa can be any amino acid, CDS does not include  
FT no start codon"  
FT /transl\_except= (pos:1582..1584, aa:Xaa)  
FT /partial  
XX  
XX US2002199217-A1.  
XX 26-DEC-2002.  
PD



QY 541 IleGlnArgGlnLeuGlnTrpLeuProGluCysLeuSer 553  
 Db 1621 ATTGAGGTCAGTGCATGCTGCGGAGTGTATCT 1659

## RESULT 2

ADAG9666  
 ID ADAG9666 standard; DNA; 1377 BP.

XX ADAG9666;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 2989.

XX Plant; bacterial infection; fungal infection; viral infection; rice;  
 KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.

XX Claim 6; SEQ ID NO 2989; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.

XX SQ Sequence 1377 BP; 290 A; 338 C; 360 G; 388 T; 0 U; 1 Other;

## Alignment Scores:

Pred. No.: 1.95e-145 Length: 1377  
 Score: 1710.00 Matches: 358  
 Percent Similarity: 78.36% Conservative: 44  
 Best Local Similarity: 69.79% Mismatches: 39  
 Query Match: 61.53% Indels: 73  
 DB: 8 Gaps: 9

US-10-051-909-36 (1-553) x ADAG9666 (1-1377)

QY 44 AsnArg---GlyAlaGlyAlaGlyGluCysLeuSerGlySerAspHisAsp-----Gly 60

Db 4 AACAGGGGCGCGCGCTGCGGCTGCGGAGCGGAGCGGAGCGGAGCGGCGGC 63

QY 61 ValLeuArgArgProLeuLeu---AsnThrGlySerTrpTrpArgMetSerSerArgGln 79

Db 64 GGATGCGGAGCGCGCTGCTGTATGCACACGGGNGCTGTACAGATGGGGTCGCGGCAG 123

QY 80 SerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98

Db 124 GGGAGCCTCACCGCGCGGGACCTCGTCATGCCATCTACCGAGTCCCACGTCCTC 183  
 QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
 Db 184 GCCTTCTCTGACAGCTCATCGTCGCGCTCGGCCCATCCAGTTCGAGTTCACGGGGGC 243  
 QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138  
 Db 244 TTCTCTCCCGCAGCGAGGACGCATCATCCGAGACCTCGACCTCACCTCTCCGAGTTC 303  
 QY 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158  
 Db 304 TCGGTGTTCGATCGCTGCTCAACGTCGCGGCCATGTTGGGGGATTCGACGTGGTCAA 363  
 QY 159 MetAlaGluTrpIleGlyArgLysGly----- 167  
 Db 364 ATGCCCGAGTACATTGGCGCAAGGKGTGAGAATTTTTTCTCTTTTGGGTACACATC 423  
 QY 168 SerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys 187  
 Db 424 TCATTGATGATTGCTGCAATTCCAAACATCATTTGGTTGGCTTGCCATCTCTTTGCAAAA 483  
 QY 188 AspAlaSerPheLeuTyrMetGlyArgLeuGluGlyPheGlyValGlyIleIleSer 207  
 Db 484 GACTCATCGTTCCTTATATGGGACCATCTCTCGAGGGGTTTGGTGTGTGTCATCTCT 543  
 QY 208 TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly 227  
 Db 544 TATACGGTGCACGTTTACATAGCAGAAATATACCTCAAAACATGAGAGGTGCTCTTGGC 603  
 QY 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247  
 Db 604 TCAGTGAATCAGTATATCTGTAACCGTTGGTATATTTGTCGATATTTGCTCGGCATGTT 663  
 QY 248 IleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267  
 Db 664 GTTCCTTGGAGGCTTCTTGCTGTAAAGGAATCTTGCTTGCACATGTTGTATACCTGGC 723  
 QY 268 LeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCys 287  
 Db 724 CTATTTCTTCAATCCAGAAATCCCAAGATGGTGGCAAGATGAACATGATGATGATTTT 783  
 QY 288 GluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAsp 307  
 Db 784 GAGACTTCTTTACAAGTTCTGAGGGGATTTGACACTGACATCAGCGCGGAAGTGAATGAT 843  
 QY 308 IleLysArgAlaValAlaSerSerLysArgThrThrIleSer---PheGlnGluLeu 326  
 Db 844 ATAAAG-----GTAGCATCTTCAAAGCAGCAGTCATATTTTAGCCGANTTTGGCTTCGTG 897  
 QY 327 AsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsn 346  
 Db 898 ACAGCGAAA----- 906  
 QY 347 LeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyVal 366  
 Db 907 -----ATTTCTTCAGGTCTC 921  
 QY 367 ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyVal 386  
 Db 922 ACAACAGTGTGCTGCTACATGTGCATTTGGTGTATCCCTCAAGGATAGCATTTCAAA 963  
 QY 387 ThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGly 406  
 Db 964 -----ATCTCTTCTGCTGGG 978  
 QY 407 MetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGln 426  
 Db 979 ATGACTCTAAGCCTCTTGCAGTTGCTGTGTATTTTCTCCTCAAGGATAGCATTTCAAA 1038  
 QY 427 AspSerAsnSerTyrTrpIleLeuThrMetIleSerLeuValGlyIleValSerPheVal 446

Db 1039 GATTCTCACATGTAACACCTTAAGTATGATCTCTGGTGGCTCTGTGGCTTTTGTGA 1098  
 Qy 447 IleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuPro 466  
 Db 1099 ATCGCCCTTCCTTCGGTATGGGTGCATTCATGATGATCAATAATGTCAGAGATCCTCCG 1158  
 Qy 467 ValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTyrLeuThrSerPhe 486  
 Db 1159 GTTAGTATCAAGAGTCTCGCAGGAAGCTTTGGCAGCGCTCGCCAACTGGCTTACATCCTTT 1218  
 Qy 487 AlaIleThrMetThrThrAsnLeuMetLeuThrTyrSerValGlyGlyThrPheLeuSer 506  
 Db 1219 GGAATAACATGACAGCAAACTTGATGCTTAGCTGGAGTGGCTGGAGGACCTTTGTGTC 1278  
 Qy 507 TyrMetValValSerAlaPheThrIleValPheValValLeuTyrValProGluThrIys 526  
 Db 1279 TACATGGTGGTGGTCTTTCACCCCTCGTGTGCTATCCCTTTGGTGGTCCAGACAAA 1338  
 Qy 527 Gly\*\*\*AsnSerArgGlyAspThrIlePheValSerLeu 539  
 Db 1339 GGAAG-AACTCTGAAAGATACAAATGGTCTTCCTCCGCTG 1376  
 RESULT 3  
 ADA69667  
 ID ADA69667 standard; DNA; 1428 BP.  
 XX  
 AC ADA69667;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Rice gene, SEQ ID 2990.  
 XX  
 KW Plant; bacterial infection; fungal infection; viral infection; rice;  
 KW gene; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 PN WO2003000898-A1.  
 XX  
 PD 03-JAN-2003.  
 XX  
 PF 22-JUN-2001; 2001WO-IB001105.  
 XX  
 PR 22-JUN-2001; 2001WO-IB001105.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 XX  
 DR WPI; 2003-175290/17.  
 XX  
 PT Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.  
 XX  
 PS Claim 6; SEQ ID NO 2990; 899pp; English.  
 XX  
 CC The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC the expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interfection. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 XX  
 SQ Sequence 1428 BP; 352 A; 315 C; 354 G; 407 T; 0 U; 0 Other;  
 Alignment Scores:

Pred. No.: 2.57e-133 Length: 1428  
 Score: 1577.00 Matches: 318  
 Percent Similarity: 85.85% Conservative: 34  
 Best Local Similarity: 77.56% Mismatches: 32  
 Query Match: 56.75% Indels: 27  
 DB: 8 Gaps: 2  
 US-10-051-909-36 (1-553) x ADA69667 (1-1428)  
 Qy 130 AspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGly 149  
 Db 277 GATCTGTACTTACACACATATGTTCTGGCTTCGTTTCGCTGTCAACGTCGGGCC 336  
 Qy 150 MetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeu 169  
 Db 337 ATGGTCGGAGCGATCGCCAGTGGGAGATGGCGAGTACATTTGGCCGGAAGGGTCTGTG 396  
 Qy 170 MetIleAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAla 189  
 Db 397 ATAATTTCGGCGGTTCCTTAACATCATTTGGTTGGCTTCCTTTGCAAAAGACGCG 456  
 Qy 190 SerPheLeuTyrMetGlyArgLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209  
 Db 457 TCATTTTATATACATGGGACGCTTCGTTGAAGGTTTGGTGTGTATATCATATACG 516  
 Qy 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229  
 Db 517 GTGCCAGTATATACATAGCAGAGATATCTCATCAGAACACAAAGAGGAGCCTTGGCTCGGTG 576  
 Qy 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePro 249  
 Db 577 AACCAAGTGTCCGTTACCATTTGGTATCTTGTGGCTATTTGTAGGCATGTTTGTCT 636  
 Qy 250 TrrArgIleLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhe 269  
 Db 637 TGGAGCTGCTTCGAGTATAGGAGCATCCCATGATCATTTGTAATACCTGCTGCTATTTC 696  
 Qy 270 PheIleProGluSerProArgTrrLeuAlaLysMetAsnLeuThrGluAspCysGluThr 289  
 Db 697 TTCATTCGGGAATCCCAAGATGGCTGGCAAAATGAAATGATGATGATTTTGAGGCT 756  
 Qy 290 SerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluValAlaAsnAspIleLys 309  
 Db 757 TCTCTACAAGTTTTGAGGGGGTTTGAGACTGATATTACTGCAGAA----- 801  
 Qy 310 ArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys 329  
 Db 802 AGACAGTAGCATCGCAACAAAGAACACACAGTCCGTTTAAAGAGTTGAACCAAAAG 861  
 Qy 330 LysTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGly 349  
 Db 862 AAATACCGCACTCCCTTACTGATAGGAACCTGGCTTCTTGACTTCAGAACTTAACTGGA 921  
 Qy 350 IleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSer 369  
 Db 922 ATAAATGGTATTCGTTTATGCAAGTAGAATCTTCAGAGATGACAGGGTTTCAACACAGT 981  
 Qy 370 AspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrrP 389  
 Db 982 GACTTGCCCATGTGCTAGGAGCAATTCAG----- 1014  
 Qy 390 LeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSerThrSerGlyMetThrLeu 409  
 Db 1015 -----ATCTACTGCTGGGATGACTCTA 1038  
 Qy 410 CysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsn 429  
 Db 1039 AGCCTTCTTCAGTATCTGTTGATTTTCTTGGAGGGTAATATTTTCACATGATTCTCAT 1098  
 Qy 430 SerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPhe 449  
 Db 1099 TCGTTCTACATCTTAAGTATGATCTCTTGGTGTGCTTGTGGCTTATATCATCATCCTTT 1158



QY 450 SerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluLeuLeuProValSerIle 469  
 Db 1159 TCCTTCGGCATGGGTGCAATTCATGGGTGCTATGATCTGAGATCCTCCAGTTAGCATC 1218  
 QY 470 LysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThr 489  
 Db 1219 AAGAGTCTGGGGAGCTTTGGACACTTGGCAACATGTTACATCTCTGGCCAAACAA 1278  
 QY 490 MetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTrpMetVal 509  
 Db 1279 ATGACAGCAAAATTTGTTACTCAGCTGGAGTCTGGAGGAAACATTTTGTCTACATGATT 1338  
 QY 510 ValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly\*\*\*Asn 529  
 Db 1339 GTGAGTGTCTTCACTCTCGTGTTCGTATATTTTGGTGGTCCAGACAAAGGAAG-AAC 1397  
 QY 530 SerArgGlyAspThrIlePheValSerLeu 539  
 Db 1398 TCTGGAGGAGATACAGTTTTCGTTCTGCTA 1427

RESULT 4

ADR59784  
 ID ADR59784 standard; CDNA; 1077 BP.

AC ADR59784;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 XX Cotton cDNA sequence, SEQ ID 565.

XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
 KW drought tolerance; plant disease resistance; galactomannan; lignin;  
 KW plant growth regulator; heat tolerance; herbicide tolerance;  
 KW homologous recombination; extreme osmotic condition tolerance;  
 KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
 KW stress resistance.

XX Gossypium hirsutum.

XX US2004181830-A1.

XX 16-SEP-2004.

XX 29-JAN-2004; 2004US-00767795.

XX 07-MAY-2001; 2001US-00849529.

XX 12-DEC-2001; 2001US-00021323.

XX (KOVA/) KOVALIC D K.  
 PA (ZHOU/) ZHOU Y.  
 PA (CAOY/) CAO Y.

XX Kovalic DK, Zhou Y, Cao Y;  
 XX WPI; 2004-667718/65.

XX New recombinant nucleic acid molecules and polypeptides from Gossypium  
 PT hirsutum useful for producing plants with improved biological  
 PT characteristics (e.g. improved plant cold or drought tolerance).  
 XX Claim 1; SEQ ID NO 565; 14pp; English.

XX The invention relates to a recombinant polynucleotide comprising any of  
 CC the 58798 Cotton plant cDNA sequences mentioned in the specification.  
 CC Also a recombinant polypeptide comprising any of the 58798 amino acid  
 CC sequences mentioned in the specification and producing a plant having an  
 CC improved property. Producing a plant having an improved property  
 CC comprises transforming a plant with a recombinant construct comprising a  
 CC promoter region functional in a plant cell operably joined to a  
 CC polynucleotide comprising a coding sequence for a polypeptide associated  
 CC with the property, and growing the transformed plant. The polypeptide is  
 CC useful for improving plant cold tolerance, manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, improving plant

CC drought tolerance, providing increased resistance to plant disease,  
 CC producing galactomannan (or lignin or plant growth regulators), improving  
 CC plant heat tolerance, improving plant tolerance to herbicides, increasing  
 CC the rate of homologous recombination in plants, improving plant tolerance  
 CC to extreme osmotic conditions or to pathogens or pests, improving yield  
 CC by modification of photosynthesis, modifying seed oil or protein yield  
 CC and/or content, improving yield by modification of carbohydrate, nitrogen  
 CC or phosphorus use and/or uptake, or improving yield by providing improved  
 CC plant growth and development under at least one stress condition. The  
 CC polynucleotide and polypeptide may also be used in recombinant DNA  
 CC constructs, in physical arrays of molecules, as plant breeding markers,  
 CC or in computer-based storage and analysis systems. The present sequence,  
 CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequences.html?DocID=20040181830. However only 6585  
 CC polynucleotide sequences were available, the remaining 52213  
 CC polynucleotides and all 58798 protein sequences were not present.

XX Sequence 1077 BP; 279 A; 186 C; 267 G; 345 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-29e-97 Length: 1077  
 Score: 1182.50 Matches: 230  
 Percent Similarity: 76.58% Conservative: 61  
 Best Local Similarity: 60.53% Mismatches: 55  
 Query Match: 42.55% Indels: 34  
 DB: 13 Gaps: 2

US-10-051-909-36 (1-553) x ADR59784 (1-1077)

QY 12 LysLeuGlyLysLysLysAsnGlnGlyLeuArgArgGluAlaValProGlyArgProAla 31  
 Db 37 AGAAAAAAGAGAGGAGGAAACATGAGTTTCAGGGATCATCTCTGAAGATGGGGAGAC--- 93  
 QY 32 SerGluLeuArgThrArgValMetGlyGlyGlySerAsnArgGlyGlyAlaGlyAlaGly 51  
 Db 93 ----- 93  
 QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71  
 Db 94 -----GTACGGAAGCCATTTCTGTCATACCGGAAGC 123  
 QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91  
 Db 124 TGGTATCGAATGGGTTCCAGACGGATCCAGTATGCTGGGCTCATCTCAA-----GTT 177  
 QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111  
 Db 178 ATTCGTGATAGTTCCATCTCTGTGTAGCTTGGCTTTTGATTGTTGCTTTGGGTCTTATC 237  
 QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131  
 Db 238 CAATTGGTTTCACTAATGGCTACTCTTCTCTACCAATCTGCAATCATCAAGGAACCTT 297  
 QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerIleSerAsnValGlyGlyMetVal 151  
 Db 298 GGGTTAACTGTATCAGAGTATTCTGTATTGGTTCTTTTCAAAATGTGGGTGCGCATGATT 357  
 QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171  
 Db 358 GGAGCAGTAGTTAGTGGTCAGATGGCAGAGTATATAGGGCGAAAGGGTCTTTTGATGATT 417  
 QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191  
 Db 418 GCTGCTATTCTTAATTAATGGATGGCTTGCTATATCTTTTGCAATGACGATCTCTCC 477  
 QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211  
 Db 478 CGGTTCGTCGCCAGCGTTGCTGGAAAGGTTTCGGTGTGGGAATAATCTCTTACACGGTGCCT 537  
 QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAonGln 231

```
Db 538 GTGTATATAGCTGAGATAGACCCGAAAAAATTGAGAGGAGCTGGTCTGTGTAATCAG 597
Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTriaArg 251
Db 598 TTGCTCTGTACAACTGGAACAATGTGTAGCTATTGCTGGGACTTTTGTGGGTGGAGG 657
Qy 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db 658 ATACTTGCATTTTAGCAGGTGTGCTGTACAATTTTGAATTCACAGGTCTATTTTCATT 717
Qy 272 ProGluSerProArgTyrLeuAlaValMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db 718 CCAGATCTCTAGATGGCTGGCAAAAATGGGAATGACCGAAGATTTCCAGACTTCTCTT 777
Qy 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLeuArgAla 311
Db 778 CAGTCTCTAGGGTTTCATCCGATATTTCTATCGAAGTCAATGAATCAAGAGTCT 837
Qy 312 ValAlaSerSerSerIleValThrThrIleSerPheGlnGluLeuAsnGlnLys-LysTy 331
Db 838 GTAGCATCAACAATAGAAAGAACACGATCCGGTTCCGAGAACTCAAAAAAAGGAGATA 897
Qy 331 rArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAs 351
Db 898 TTGGTTTCATTCATGTTGGAAATGGGCTTCTATGTTGCAACAACCTTAGTGGCATTAA 957
Qy 351 nGlyValLeuPheTyrAlaSerSerIlePhePheIleAlaGlyValThrAsnSerAspLe 371
Db 958 TGGTGTATATCTATTCCGCTACCATATTCGAAACTGCTGGAATTAATCAAGCAATAT 1017
Qy 371 uAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeu 390
Db 1018 AGCTACCTTTGGAGTTGGCTTCATTCAGTCACTGCTACTGCTTAACCACTGGTGTG 1075

RESULT 5
AAC36319
ID AAC36319 standard; DNA; 1741 BP.
AC AC AAC36319;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13356.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN
PN EP1033405-A2.
XX
PD
PD 06-SEP-2000.
XX
PF
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 16-APR-1999; 99US-0128714P.
XX 19-APR-1999; 99US-0129845P.
XX 21-APR-1999; 99US-0130077P.
XX 23-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142903P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 19-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
```

PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	22-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	22-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	22-JUL-1999;	99US-0145276P.	PR	23-OCT-1999;	99US-0161404P.
PR	22-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147326P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	08-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149475P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	23-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	28-SEP-1999;	99US-0156596P.			
PR	28-SEP-1999;	99US-0157117P.			
PR	04-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			

Alignment Scores:  
 Pred. No.: 2,68e-76 Length: 1741  
 Score: 951.50 Matches: 181  
 Percent Similarity: 66.12% Conservative: 102  
 Best Local Similarity: 42.29% Mismatches: 140  
 Query Match: 34.24% Indels: 5  
 DB: 3 Gaps: 2

US-10-051-909-36 (1-553) x AAC36319 (1-1741)

QY	100	PheLeuCyThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe	119
DB	287	TATCTCAGCACATTCGTCGGTCTGTGTCCTTTTCGGTTCGTTCTTGGGGATAC	346
QY	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer	139
DB	347	TCCTTACCTGCTCAGGCTGCAATTAGGAATGACCTTTCATTGACTATAGTAGTGTTC	406
QY	140	AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet	159
DB	407	CTCTTTGGTTCCTTACTAATCTTTGGCGCAATGATCGGTGCTATAACAAGCGGCTATA	466
QY	160	AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179
DB	467	GCTGATTAGTTGGAGAAGAGGGGCGATGAGAGTTCTCTGCAATTTGTGTAGTCGG	526
QY	180	TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu	199
DB	527	TGGCTAGCAATCATCTTCCCAAGGAGTAGTGGCTCTGGACCTCGGAAGACTGGCAACG	586
QY	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro	219
DB	587	GGATATGGAATGGGACCATTTTCTATGTGGTCCCAATCTTTATAGCAGAAATTCACCT	646
QY	220	GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe	239
DB	647	AAACTTTTCAGAGGGGCTTTAACCACTGACACAGATTCGTATCTGCATGGAGTGTG	706
QY	240	LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu	259
DB	707	GTTTCTTCATCATAGGCACACTAGTAGCTGGAGAGTCTTGGCATTAAATAGGAATCATC	766
QY	260	ProCysThrMetLeuIleProGlyLeuPheIleProGlySerProArgTyrPheLeuAla	279
DB	767	CCATGGCTGCT	826
QY	280	LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr	299
DB	827	AAAGTGGGGCGTGATACGGAGTTTGAAGCTGCATGAGGAGCTCGTGGGAAGAGGCT	886
QY	300	AspIleThrThrGluValAlaAsnAspIleLysArgAlaValAlaSerSerLysArgThr	319
DB	887	GATATTTCCGAGGAGGAGCAGAGATCCAGGATTATATCGAAACTCTGGAAGAGCTACC	946
QY	320	ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIle	339



QY	153	AlaIleAAsSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla	172	PR	19-APR-1999;	99US-0130077P.
Db	418	GCAATAGCCAGTGGTCAGATAGCCGAATGATTTGGACCGAAGGGTCTTTAATCACTGCT	477	PR	21-APR-1999;	99US-0130449P.
QY	173	AlaIle-ProAAsIleIleGlyTrpLeuAlaIleSerPheAlaIleAspAlaSerPheLe	192	PR	23-APR-1999;	99US-0130510P.
Db	478	GCAATTCCTTAAATATATTGGATGGCTTGCTATATCTTTTGCAGAGATTCCTCTTTCT	537	PR	28-APR-1999;	99US-01310891P.
QY	192	uTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProVa	212	PR	30-APR-1999;	99US-0131449P.
Db	538	TTACATGGGAGGAGCTGTGCAAGTCTCTGTCGCGCATATCTTATACGGTGGCTGT	597	PR	30-APR-1999;	99US-0132048P.
QY	212	lTyrIleAlaGluIleSerProGlnAAsMetArgGlyAlaLeuGlySerValAAsnGlnLe	232	PR	04-MAY-1999;	99US-0132407P.
Db	598	ATATATAGCTGAGATAGCACTCAGAAATTTGAGGGGCAGCTTGGGTTTCAGCGAATCAGCT	657	PR	05-MAY-1999;	99US-0132484P.
QY	232	uSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProThrArgLe	252	PR	06-MAY-1999;	99US-0132485P.
Db	658	ATCTGTCACTCTCGGAATAATGCTGGCCTATCTACTTCGACTTTTGTGAGTGCCGGAT	717	PR	07-MAY-1999;	99US-0132486P.
QY	252	uLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIlePr	272	PR	07-MAY-1999;	99US-0132487P.
Db	718	ACTCGCGGTTTAGGAATACCTCTGTGACAAATTTGATACCTGCTCTATTTTCATTC	777	PR	11-MAY-1999;	99US-0132488P.
QY	272	oGluSerProArgTrpLeuAlaLysMetAAsnLeuThrGluAAspCysGluThrSerLeuGl	292	PR	14-MAY-1999;	99US-0132489P.
Db	778	AGAATCTCTCGATGGTGGCAAACTGGGTATGACAGAGAATTCGAACTCTTTTGCA	837	PR	14-MAY-1999;	99US-0132491P.
QY	292	nValLeuArgGlyPheGluThrAspIleThrGluValAAsnAspIleLysArgAlaVa	312	PR	14-MAY-1999;	99US-0132492P.
Db	838	AGTTCTCAGGTGCTCCGATGCTGATATTTCTATTGAAGTGAATGAATCAACAGATCTGT	897	PR	18-MAY-1999;	99US-0132493P.
QY	312	lAlaSerSerSerLys-ArgThrThrIleSerPheGlnGluLeuAAsnGlnLysLysTyrA	332	PR	18-MAY-1999;	99US-0132494P.
Db	898	CTCATCAACCACTAGCAAGAACTACGATTCGGTTTG--CACATCTCAAGAAAGATATT	955	PR	01-JUN-1999;	99US-0132495P.
QY	332	rgThrProLeuLeuLeuGlyIleGlyLeuLeuVal	343	PR	03-JUN-1999;	99US-0132496P.
Db	956	GGTCTCATTTGATGGTTGGAATTCGCTCACTTATG	990	PR	04-JUN-1999;	99US-0132497P.
RESULT 7						
AC	AAC39099					
AC	ID AAC39099 standard; DNA; 1683 BP.					
XX						
XX	AAC39099;					
XX						
DT	17-OCT-2000 (first entry)					
XX						
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 23364.					
XX						
KW	Hybridisation assay; genetic mapping; gene expression control;					
KW	protein identification; signal transduction pathway; metabolic pathway;					
KW	promoter; termination sequence; ss.					
XX						
OS	Arabidopsis thaliana.					
XX						
PN	EP1033405-A2.					
XX						
PD	06-SEP-2000.					
XX						
PF	25-FEB-2000; 2000EP-00301439.					
XX						
PR	25-FEB-1999; 99US-0121825P.					
PR	05-MAR-1999; 99US-0123180P.					
PR	09-MAR-1999; 99US-0123548P.					
PR	23-MAR-1999; 99US-0125788P.					
PR	25-MAR-1999; 99US-0126264P.					
PR	28-MAR-1999; 99US-0126785P.					
PR	01-APR-1999; 99US-0127462P.					
PR	06-APR-1999; 99US-0128234P.					
PR	08-APR-1999; 99US-0128714P.					
PR	16-APR-1999; 99US-0129845P.					

PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144864P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147436P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			

Alignment Scores:

Pred. No.:

1.36e-75

Length:

1683

Score:

943.50

Matches:

180

Percent Similarity:

65.89%

Conservative:

102

Best Local Similarity:

42.06%

Mismatches:

141

Query Match:

33.95%

Indels:

5

DB:

3

Gaps:

2

US-10-051-909-36 (1-553) x AAC39099 (1-1683)

Qy	100	PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe	119
Db	207	TATCTCAGCACATTCGTCGGTCTGTGGTCTTTTCGCTCGTCTTCTTCTCGGGATC	266
Qy	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer	139
Db	267	TCTTCACCTGCTCAGGTCGCAATTAGGAATGACCTTTTCATTGACTATAGCTGAGTTTCA	326
Qy	140	AlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMet	159
Db	327	CTCTTTGGTCTTTTACTAACTTTTGGCGCAATGATCGTCTATACACGCGGCTATA	386
Qy	160	AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179
Db	387	GCTGATTTAGTTGCAAGAAAGGGCGCATGAGACTTTCTCTGCAATTTTGTAGTCGGG	446
Qy	180	TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu	199
Db	447	TGGCTAGCAATCTTTTCCAGGGAGTAGTGCTCTGGACCTCGGAGACTGGCAACG	506
Qy	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro	219
Db	507	GGATATGGAATGGAGACATTTTCTTATGTCGCCAATCTTTATAGCAGAAATGCACT	566
Qy	220	GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe	239
Db	567	AAAACTTTCAGAGGGGCTCTAACCACTGAACCACTGATCTGATCTGCACTGGAGTTCG	626
Qy	240	LeuAlaTyrLeuLeuGlyMetPheIleProTyrPargLeuLeuAlaValIleGlyAlaLeu	259
Db	627	GTTCCTTTCATCATTAGGCACACTAGTAGCCTGGAGACTTTCGGCATTAAATAGGAATC	686
Qy	260	ProCysThrMetLeuIleProGlyPhePheIleProGluSerProArgTyrTrpLeuAla	279
Db	687	CCATGCGCTGCTCTCTTCTTGGCTCTTTTATATCCCTGAGTCTCCAGATGGCTGGCA	746
Qy	280	LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr	299

```
Db 747 AAGTGGGGCTGATACGGAGTTTGAAGCTGCACCTGAGGAAGCTCCGTGGGAGAGGCT 806
Qy 300 AspleThrThrGluValAenAspIleGlyeArgAlaValAlaSerSerLysArgThr 319
Db 807 GATATTTCCGAGGAGGAGCAGACGATCCAGGATATATCGAAACTCTGGAAGGCTACCG 866
Qy 320 ThrIleSerPheGlnGluLeuAenGlnLysValysTyrArgThrProLeuLeuGlyIle 339
Db 867 AAGCCAAAGATGCTGGATTTGTTTCAGAGGAGATACATACGCTCTGTCTTATAGCTTTC 926
Qy 340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer 359
Db 927 GGGTGTAGTGTTCACGAGTTTGGAGGATCAACGGAATATGTTTCTACACAGCTCG 986
Qy 360 IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379
Db 987 ATATTTGAGCAAGCAGGCTTTTCCC---ACAAGACTTGGGATGATAATATATGCTGTCTT 1043
Qy 380 GlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeu 399
Db 1044 CAGGTGTAATCACTGCGCTTAATGACCCGATAGTTGACAGAGCGGGAAGAAACCATTTG 1103
Qy 400 LeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePhe 419
Db 1104 CTACTGGTTTCTCAACAGGTTAGTAGTGGTGTGTTGATAGCAGCGGTTTCTTCTAT 1163
Qy 420 LeuLysAspAenIleSerGlnAspSerAsnSerTyrTrpIleLeuThrMetIleSerLeu 439
Db 1164 CTCAAGGTTCCAC-----GACATGCGCGCACGAAGCAGTCCCGCTGCTGT 1211
Qy 440 ValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeu 459
Db 1212 GTTGGTATATGTTGATACATAGGATCGTTTTCACAGGAAATGGGAGCAATGCCGTGGTG 1271
Qy 460 MetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeu 479
Db 1272 GTCATGTCTGAGATATTTCCATAAACATAAAGGAGTACGAGGAGCATGGGACGCTG 1331
Qy 480 AlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSer 499
Db 1332 GTGAAGTGGTTTGGAGCGTGGCTGTTTCTTACACTTTCAACTTCCTCATGCTCGGAGC 1391
Qy 500 ValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVal 519
Db 1392 TCTTACGGAACTTTCTCTCATTTAGCTGCCATCAACGCGCTGGCCATTGTTCTGTCATT 1451
Qy 520 LeuTrpValProGluThrLysGly 527
Db 1452 GCCATCGTGCCTGAGACAAAGGG 1475

RESULT 8
AAC45857
ID AAC45857 standard; DNA; 1395 BP.
XX
AC AAC45857;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48015.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 03-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 08-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143542P.
```

PR	14-JUL-1999;	99US-0143C24P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	17-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	18-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	18-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144684P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	22-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	30-AUG-1999;	99US-0151080P.			
PR	31-AUG-1999;	99US-0151303P.			
PR	01-SEP-1999;	99US-0151438P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
			Alignment Scores:		
			Pred. No.:	9.31e-75	1395
			Score:	933.00	Matches: 180
			Percent Similarity:	65.73%	Conservative: 102
			Best Local Similarity:	41.96%	Mismatches: 141
			Query Match:	33.57%	Indels: 6
			DB:	3	Gaps: 3
			US-10-051-909-36 (1-553) x AAC45857 (1-1395)		
QY	100	PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119	QY	100	PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119
Db	85	TATCTCAGCACATTCGTTGGGTTCTTTTGGTTCGTTCTTTGGTGGGATAC 144	Db	85	TATCTCAGCACATTCGTTGGGTTCTTTTGGTTCGTTCTTTGGTGGGATAC 144
QY	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139	QY	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139
Db	145	TCTTCACCTGCTCAGGCTGCAATAGGAATGACCTTTCATTGACTATAGCTGATTTTCA 204	Db	145	TCTTCACCTGCTCAGGCTGCAATAGGAATGACCTTTCATTGACTATAGCTGATTTTCA 204
QY	140	AlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMet 159	QY	140	AlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMet 159
Db	205	CTCTTTGGTCTTTTACTTAACCTTTTGGCGCAATGATCGTGTCTATACAGCGGCTATA 264	Db	205	CTCTTTGGTCTTTTACTTAACCTTTTGGCGCAATGATCGTGTCTATACAGCGGCTATA 264
QY	160	AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179	QY	160	AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179
Db	265	GCTGATTTAGTTGGAAGAGGGGGGATGAGAGTTTCTCTGCAATTTGTTGTAGTCGGG 324	Db	265	GCTGATTTAGTTGGAAGAGGGGGGATGAGAGTTTCTCTGCAATTTGTTGTAGTCGGG 324
QY	180	TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuGlu 199	QY	180	TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuGlu 199
Db	325	TGGCTAGCAATCATCTTTGCAAGGAGTAGTGGCTCTGGACCTTGGAAAGCTGCAACG 384	Db	325	TGGCTAGCAATCATCTTTGCAAGGAGTAGTGGCTCTGGACCTTGGAAAGCTGCAACG 384
QY	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219	QY	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219
Db	385	GGATATGGAATGGAGCATTTCTTATGTGTGCCAATCTTTATAGCAGAAATTCACCT 444	Db	385	GGATATGGAATGGAGCATTTCTTATGTGTGCCAATCTTTATAGCAGAAATTCACCT 444
QY	220	GlnAsnMetArgGlyAlaLeuGlySerValAsn--GlnLeuSerValThrPheGlyIle 238	QY	220	GlnAsnMetArgGlyAlaLeuGlySerValAsn--GlnLeuSerValThrPheGlyIle 238
Db	445	AAAACCTTCAGAGGGCTCTTAACACACACTGAACACAGACATTCGTGCACTGGAGTG 504	Db	445	AAAACCTTCAGAGGGCTCTTAACACACACTGAACACAGACATTCGTGCACTGGAGTG 504
QY	239	PheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAla 258	QY	239	PheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAla 258





```
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 08-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 18-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145193P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0145955P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149436P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 05-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 9,97e-74
Score: 923.00
Percent Similarity: 65.58%
Best Local Similarity: 42.09%
Query Match: 33.21%
DB: 3
US-10-051-909-36 (1-553) x AAC45298 (1-1685)

QY 100 PheLeuCysThrLeuIleValAlaLeuGlyPProIleGlnPheGlyPheThrSerGlyPhe 119
Db 207 TAICTCAGACATTCTGGGTCTGGTCTCTTTGGCTTCGTTCTTTGGTTCGCGGATAC 266
QY 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139
Db 267 TCTTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTCACTAGCTATAGCTAGTTCATCA 326
QY 140 AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet 159
Db 327 CTCTTTGGTTCTTTTACTAACTTTTGGCGCAATGATCGGTGCTTATAACAAGCGGCTATA 386
QY 160 AlaGluTyrlIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179
Db 387 GCTGATTTAGTTGGAAGAAAGGGGCGATGAGAGTTCTCTGCACTTTTGTGTAGTCGGG 446
QY 180 TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyMetGlyArgLeuLeuGlu 199
Db 447 TGGCTAGCAATCATCTTTCCCAAGGAGTAGTGGCTCTGGACCTTTGGAAGACTGGCAACG 506
```

Qy 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219  
 Db 507 GGATATGGAATGGAGCATTTTCTATGTGTGCAATCTTTATAGCAGAAATTCACAC 566  
 Qy 220 GlnAenMetArgGlyAlaLeuGlySerValAenGlnLeu---Ser-ValThrPheGlyIle 238  
 Db 567 AAAATTTTCAGAGGGCTCTAACACACACTGACAGACAGATCTGTATCTGCACCTGGAGT 626  
 Qy 238 ePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAl 258  
 Db 627 GTCGGTTTCTTCATCATAGGCACACTAGTCAGCTGGAGAGTCTTGGCATTATAGCAAT 686  
 Qy 258 aLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLe 278  
 Db 687 CATCCCATGGCTGCTCTCTTCCTTGGCTCTTTTATCCCTGAGTCTCCAGATGGCT 746  
 Qy 278 uAlaLeuMetAenLeuThrGluAenCysGluThrSerLeuGlnValLeuAenGlyPheG 298  
 Db 747 GGCMAAGTGGGGCTGATACGGAGTTTGAAGCTGCACTAGGAGAGTCTCGTGGGAAGAA 806  
 Qy 298 uThrAspIleThrThrGluValAenAenPheIleLysArgAlaValAalaSerSerLysAr 318  
 Db 807 GGCTGATATTTCGAGGAGGCGCAGCAGAGATCCAGGATTATATCGAAACTCTGGAAAGGCT 866  
 Qy 318 gThrThrIleSerPheGlnGluLeuAenGlnLysLysTyrArgThrProLeuLeuLeuG 338  
 Db 867 ACCGAAACCAAGATGCTGGATTTGTTTTCAGAGGATACATACGCTCTGTCTTATAGC 926  
 Qy 338 yIleGlyLeuLeuValLeuAenLeuSerGlyLeuAenGlyValLeuPheTyrAlaLe 358  
 Db 927 TTTGGGTTGATGTGTTTTCAGCAGTTTGGAGGAAATCAACGAATATGTTTCTACACAAG 986  
 Qy 358 rSerIlePheLysAlaAlaGlyValThrAenSerAspLeuAlaThrCysSerLeuGlyAl 378  
 Db 987 CTCGATATTGAGCAACAGCGTTTTCCTCC---ACAAGACTTGGGATGATAATATATGCTGT 1043  
 Qy 378 aIleGlnValLeuAlaThrGlyValThrThrTrpLeuAenAspArgAlaGlyArgAl 398  
 Db 1044 TCTTCAGTGTGTAATCATCTGCTGCTTAATGCACCGCATAGTTGACAGAGCCGGAAGAAACC 1103  
 Qy 398 eLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPh 418  
 Db 1104 ATTGCTACTGTTTCTCCACAGGTTAGTGATAGGCTGTTTGTATGACAGCGGTTTCTTT 1163  
 Qy 418 ePheLeuLysAspAenIleSerGlnAenSerAsnSerTyrTyrIleLeuThrMetIleSe 438  
 Db 1164 CTATCTCAAGTTTCAC-----GACATGCGCGCAGCAGCAGTCCCGCTCTGGC 1211  
 Qy 438 rLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTr 458  
 Db 1212 TGTGTTGGTATATGTTGATCATATAGGATCGTTTTCAGCAGGAATGGGCAATGCGGTG 1271  
 Qy 458 pLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaTh 478  
 Db 1272 GGTGGTCTGCTGAGATATTTCCCAATAACATAAAGGAGTAGCAGGAGCATGGCGAC 1331  
 Qy 478 rLeuAlaAenTrpLeuThrSerPheAlaIleThrMetThrThrAenLeuMetLeuThrTr 498  
 Db 1332 GCTGGTGAATCTGTTGGAGCGTGGTGTCTTCTTACATTTCACTTCTCATGCTCTG 1391  
 Qy 498 pSerValGlyIleThrPheLeuSerTyrMetValSerAlaPheThrIleValPheVa 518  
 Db 1392 GAGCTCTTACGGAATTTCTCTCATTTACGCTGCGCATCAACGCGCTGGCCATTGTCTTCGT 1451  
 Qy 518 lValLeuTrpValProGluThrLysGly 527  
 Db 1452 CATTGCCATCTGCTGCTGAGACAAAGG 1479

RESULT 10  
 AAC45290  
 ID AAC45290 standard; DNA; 1743 BP.  
 XX  
 AC AAC45290;

XX 18-OCT-2000 (first entry)  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 45977.  
 DE Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 XX EP1033405-A2.  
 XX 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-00301439.  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 29-MAR-1999; 99US-0126264P.  
 PR 01-APR-1999; 99US-0126785P.  
 PR 06-APR-1999; 99US-0127462P.  
 PR 08-APR-1999; 99US-0128234P.  
 PR 16-APR-1999; 99US-0128714P.  
 PR 19-APR-1999; 99US-0129845P.  
 PR 21-APR-1999; 99US-0130077P.  
 PR 23-APR-1999; 99US-0130449P.  
 PR 28-APR-1999; 99US-0130510P.  
 PR 30-APR-1999; 99US-0130891P.  
 PR 04-MAY-1999; 99US-0131449P.  
 PR 06-MAY-1999; 99US-0132048P.  
 PR 07-MAY-1999; 99US-0132407P.  
 PR 11-MAY-1999; 99US-0132484P.  
 PR 14-MAY-1999; 99US-0132485P.  
 PR 14-MAY-1999; 99US-0132486P.  
 PR 14-MAY-1999; 99US-0132487P.  
 PR 18-MAY-1999; 99US-0132863P.  
 PR 20-MAY-1999; 99US-0134941P.  
 PR 21-MAY-1999; 99US-0135124P.  
 PR 24-MAY-1999; 99US-0135353P.  
 PR 25-MAY-1999; 99US-0135629P.  
 PR 27-MAY-1999; 99US-0136021P.  
 PR 28-MAY-1999; 99US-0136392P.  
 PR 03-JUN-1999; 99US-0136782P.  
 PR 04-JUN-1999; 99US-0137222P.  
 PR 07-JUN-1999; 99US-0137528P.  
 PR 08-JUN-1999; 99US-0137502P.  
 PR 10-JUN-1999; 99US-0137724P.  
 PR 10-JUN-1999; 99US-0138094P.  
 PR 14-JUN-1999; 99US-0138540P.  
 PR 16-JUN-1999; 99US-0138847P.  
 PR 18-JUN-1999; 99US-0139119P.  
 PR 18-JUN-1999; 99US-0139452P.  
 PR 18-JUN-1999; 99US-0139453P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.







PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145513P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 06-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149829P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
 Pred. No.: 4.59e-73  
 Score: 916.00  
 Percent Similarity: 61.99%  
 Best Local Similarity: 41.25%  
 Length: 1755  
 Matches: 191  
 Conservative: 96  
 Mismatches: 165

production of stress-regulated genes in response to abiotic stress with increased tolerance to abiotic stress. The present sequence is that of an *Arabidopsis thaliana* stress regulated gene (ABZ12196-ABZ17574) used



[illegible][illegible]

DD ICCCICGGI I I A I C A I I I C A I I A C A G----- 1035

Qy	432	lyrlleuunrmetreserleuvalglyllevalserphevalreunrphesulrnc	432
Db	1060	-----GTGTACACCGGATCATTTTCATTA	1083
Qy	452	GlyMetGlyAlaIleProTtpLeuMetMetSerGluIleLeuProValSerIleLysSer	471
Db	1084	GGGATGGGTGGGATTCCTGGGTCATATGTCCAGAGATATTTCCATATAGACATATAAGGA	1143
Qy	472	LeuGlyGlySerIleAlaThrLeuAlaAsnTtpLeuThrSerPheAlaIleThrMetThr	491
Db	1144	TCAGCGGAAGCCTCGTACTGTGTGTAGCTGGGTCGGATCATGGATTATATCTTTTACA	1203
Qy	492	ThrAsnLeuMetLeuThrTtpSerValGlyGlyThrPheLeuSerTyrrMetValValSer	511
Db	1204	TTTAACATTTCTAATGAATTGGAATCCGGCAGGAACGTTTATGTGTTTCGCTACTGTTTGT	1263
Qy	512	AlaPheThrIleValPheValValLeuTtpValProGluThrLysGly***AsnSerArg	531
Db	1264	GGGGCTACTGTTATTTTGTAGCAAAACTCGTACCAGAAACCAAGGTCG-CACACTTGA	1322
Qy	532	GlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTtpLeu	548
Db	1323	GGAAATCCCAATATTC-----GATTGGTTTA	1346
RESULT 14			
ADM47939			
ID	ADM47939	standard; DNA; 1768 BP.	
XX			
AC	ADM47939;		
XX			
XX	03-JUN-2004	(first entry)	
XX			
DE		Polynucleotide sequence #357 useful in producing transgenic plants.	
XX			
KW		Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;	
KW		osmotic stress; sugar transport; cell cycle pathway; plant height;	
KW		carbohydrate transport; crop productivity; plant growth;	
KW		stress resistance; disease resistance; insect resistance; heat tolerance;	
KW		nitrogen assimilation; water stress tolerance;	
KW		photosynthetic carbon fixation; virus resistance; gene therapy; gene; do.	
OS		Glycine max.	
XX			
XX		US2003233670-A1.	
XX			
PD		18-DEC-2003.	
XX			
PF		04-DEC-2002; 2002US-00310154.	
XX			
PR		04-DEC-2001; 2001US-0337358P.	
XX			
PA		(EDGE/) EDGERTON M D.	
PA		(CHOM/) CHOMET P S.	
PA		(LACC/) LACCETTI L B.	
XX			
PI		Edgerton MD, Chomet PS, Laccetti LB;	
XX			
DR		WPI; 2004-061374/06.	

## RESULT 14

ADN47939  
ID ADN47939 standard: DNA: 1768 BP.

XX  
AC ADM47939:

XX DT 03-JUN-2004 (first entry)

XX DE Polymicteotide semence #357 useful in producing transgenic plants.

XX Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;  
 KW osmotic stress; sugar transport; cell cycle pathway; plant height;  
 KW carbohydrate transport; crop productivity; plant growth;  
 KW stress resistance; disease resistance; insect resistance; heat tolerance;  
 KW nitrogen assimilation; water stress tolerance;  
 KW photosynthetic carbon fixation; virus resistance; gene therapy; gene; do.

OS Glycine max.

PN US2003233670-A1.

XX  
PD  
18-DEC-2003

04-DEC-2002: 2002US-00310154.

XX  
PR 04-DEC-2001: 2001US-0337358P.

XX PA (EDGE/) EDGERTON M D.

PA (CHOM//) CHOMEI F S.  
PA (LACC/) LACCEITI L B.

PT Edgerton MD. Chomet PS. Laccetti LB:

XX  
DR WPI: 2004-061374/06.

Fri Apr 15 06:53:45 2005

DR P-PSDB; ADM48307.  
 CC New polynucleotide, useful for manipulating plant protein quality.  
 PT Improving plant growth, yield and crop productivity or grain composition  
 PT or producing plants with improved properties.  
 XX  
 XX Claim 1; SEQ ID NO 357; 144pp; English.  
 PS  
 CC The present invention relates to polynucleotide sequences, and the  
 CC proteins they encode. The sequences are isolated from a variety of  
 CC organisms such as plants (e.g. maize, rice, sorghum, chole cress,  
 CC soybean, and wheat), cyanobacteria, bacteria, and other fungi. The  
 CC polynucleotide and polypeptide sequences of the invention are useful in  
 CC the production of transgenic plants that have improved properties. Also  
 CC disclosed are methods of producing fertile transgenic plants, preferably  
 CC maize, with desired phenotypes. The polynucleotide and polypeptide  
 CC sequences are useful for improving plants by providing protection against  
 CC osmotic stress, improving altering sugar transport and/or metabolism,  
 CC modifying the cell cycle pathway, reducing plant height, modifying plant  
 CC carbohydrate transport, improving crop productivity, improving plant  
 CC growth and stress resistance, improving disease resistance, improving  
 CC insect resistance, improving cold or heat tolerance, improving nitrogen  
 CC assimilation, improving stalk strength, improving water stress tolerance,  
 CC improving photosynthetic carbon fixation, improving biotic and abiotic  
 CC stress resistance, improving resistance to oxidative stress, providing  
 CC increased vigour, reducing senescence, and conferring stress resistance.  
 CC The present sequence represents a polynucleotide sequence of the  
 CC invention. Note: The sequence data for this patent is not provided in the  
 CC printed specification but is obtained in electronic format from the USPTO  
 CC website at seqdata.uspto.gov.  
 XX  
 XX Sequence 1768 BP; 516 A; 300 C; 384 G; 568 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5.73e-71 Length: 1768  
 Score: 893.00 Matches: 191  
 Percent Similarity: 56.25% Conservative: 97  
 Best Local Similarity: 37.30% Mismatches: 173  
 Query Match: 32.13% Indels: 51  
 DB: 12 Gaps: 4

US-10-051-909-36 (1-553) x ADM47939 (1-1768)

QY 43 SerAenArgGly-GlyAlaGlyAlaGlyGluSerGlySerAspHisAspGlyValle 62  
 DB 119 TCNACCAGAGATGCTGCTGCTGCTTGTAAAGAAATGGATCAGAT- 164  
 QY 62 uArgArgProLeuLeuAenThrGlySerTyrTyrArgMetSerSerArgGlnSerPh 82  
 DB 165 -----AAGAGTGT 172  
 QY 82 eAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCy 102  
 DB 173 GAAATATGATCCATGAGTGGTGTGTTG-----CTTAG 205  
 QY 102 sThrLeuLeuValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerPr 122  
 DB 206 CACACTGTGTGCTGTGTGTGTGCTTTCACATTTGGAATGCTGTGTGGCTATTGAGCACC 265  
 QY 122 oThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheG 142  
 DB 266 CACTCAAGCGGTATCAGGGCGAGATCTTAATCTCTCTGCTGATGTTTCCAGTTTGG 325  
 QY 142 ySerLeuSerAsnValGlyGlyMetValGlyValAlaIleAlaSerGlyGlnMetAlaGluTy 162  
 DB 326 TTCAATAGTACCATTTGGTGGCAATGCTTGGAGCTATACACCGCGCAGGATTACAGATT 385  
 QY 162 rIleGlyArgGlyGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTTPLeuAl 182  
 DB 386 CATGGCCGCAAGAGGGCAATGAGGATTTCACAGATTGTTGATTACAGATGATGATG 445  
 QY 182 aIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyGlyPheG 202  
 DB 182 alileSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyGlyPheG 202

DB 446 TGTCTCTCTCTAAGGGTCTTACTCCCTGACTTCGGAAGATTTTTCACAGCTTACGG 505  
 QY 202 yValGlyIleIleSerTyrThrValProValTyrIleAlaGluLeuSerProGlnAsnMe 222  
 DB 506 CATTCGGATATCTATCATACGCTGCTGTGTATATATAGCAGAAATAGCAACCAATCT 565  
 QY 222 tArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTy 242  
 DB 566 TCGAGGAGGACTAGCAACAATCAGCTTTTGTATTACTGGAGGCTCAGTCTCAT 625  
 QY 242 rLeuLeuGlyMetPheIleProIleArgLeuLeuAlaValIleGlyAlaLeuProCysTh 262  
 DB 626 CTATATTAGGAAGTGTATATAAATTGGAGAGAACTTGCATTTAGCGGGGTAGTGCCTTCAT 685  
 QY 262 rMetLeuIleProGlyLeuPhePheIleProGluSerProArgTTPleuAlaLysMetAs 282  
 DB 686 TTGCTTGTGCTGTTGTTGCTTTTATCCCTGAGTCCCAAGATGGCTTGGCTAAGTTGG 745  
 QY 282 nLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleTh 302  
 DB 746 CCGTGGAAAGAAATTTCAACTAGCTTTAAGTAGACTTCGGGTAAAGATGCTGATATTC 805  
 QY 302 rThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrIleSe 322  
 DB 806 TGAATGAAGCTGCTGAAATTTCTGGATATTATGAAACTCTTCAAGTCTTCTCAAGACTAA 865  
 QY 322 rPheGlnGluLeuAsnGlnLysIleLysTyrArgThrProLeuLeuGlyIleGlyLeu 342  
 DB 866 GCTGTTGGATTTGTTCCAAAGCAAAATATGTGCACCTCTGTATTGCGGTCGCTTAAT 925  
 QY 342 uValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLy 362  
 DB 926 GGCATGTCACAAATCTGTTGGAATTAATGGCATGATTTACACAGCTGAGATTTTGT 985  
 QY 362 sAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnVal 382  
 DB 986 AGCAGCTGGACTTCTTCAGAGAAAGCTGCTACATAGATGCTGTATACAGATTC 1045  
 QY 382 uAlaThrGlyValThrThrTrileuLeuAspArgAlaGlyArgGlyIleLeuIle 402  
 DB 1046 ATTTACATATTATGGAGCCATTTTGTATGACCAAGTCTGGAAGACCTCTTGTAAATGT 1105  
 QY 402 eSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPheLeuLysAs 422  
 DB 1106 TTCGCGCAGCTGGGACATCTTGTGCTTGTGTGCTGCTTGTGCTTCTTCTTCAAGGA 1165  
 QY 422 pAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGly 442  
 DB 1166 C-----CAAGCTTATGCTGAGTGGGTACCTATATTAGTATTGCTGCGCT 1213  
 QY 442 eValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProIlePheMetMetSe 462  
 DB 1214 TCTGATCTACATAGCAGCANTTTCAATTGGCTCGGATCAGTTCATGGGTGATAATGTC 1273  
 QY 462 rGluIleLeuProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAsnTr 482  
 DB 1274 TGAGATATTTCCATACATTTTGAAGGGAGCTGCTGGAAGCTTGGTGGTGGTGGT 1333  
 QY 482 pLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGly 502  
 DB 1334 GCTAGAGCTTGGGTAGTTCATATATCTTCACTTCTTATGAGCTGAGTGTCTCTCG 1393  
 QY 502 yThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpVa 522  
 DB 1394 TACATTTGTTTGTATGCTGATGTTCCCTTAACTATTCTATTGTTAGCAAAATAGT 1453  
 QY 522 lProGluThrIlyGly\*\*\*AsnSerArgGlyAspThrIlePheValSerLeuSerIleG 542  
 DB 1454 CCAGAAACCAAGGA----- 1469  
 QY 542 nArgGlnLeuGlnTrpLeuProGluCysLeuSer 553  
 DB 1470 -AAACTTTGGAGAGATCCAGGCTTGTATTAGT 1502

## RESULT 15

ABZ13486  
ID ABZ13486 standard; DNA; 1407 BP.

XX AC ABZ13486;  
XX DT 21-JAN-2003 (first entry)  
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1291.  
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX OS Arabidopsis thaliana.  
XX PN W0200216655-A2.  
XX PD 28-FEB-2002.  
XX XX 24-AUG-2001; 2001WO-US026685.  
XX PR 24-AUG-2000; 2000US-0227866P.  
XX PR 26-JAN-2001; 2001US-0264647P.  
XX PR 22-JUN-2001; 2001US-0300111P.  
XX XX (SCRI ) SCRIPPS RES INST.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX PI Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX Claim 144; SEQ ID NO 1291; 577pp + Sequence Listing; English.  
XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX SQ Sequence 1407 BP; 350 A; 261 C; 331 G; 465 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1.05e-69	Length:	1407
Score:	877.50	Matches:	179
Percent Similarity:	59.06%	Conservative:	98
Best Local Similarity:	38.17%	Mismatches:	166
Query Match:	31.58%	Indels:	26
DB:	6	Gaps:	5

US-10-051-909-36 (1-553) x ABZ13486 (1-1407)

QY	66	LeuLeuAsnThrGlySerTptTyArg-MetSerSerArgGlnSerSerPheAlaProG1	85
DB	12	CTCAAGAACTTGGAGCTGGTTGCTACTGAGGAGAACCAACGAC-----	60
QY	85	yThrSerSerMetAlaValLeuArgGluSerHisValSerAla-----PheLeuCysTh	103
DB	61	-----ATCAACGAATTCGTATCATCTGCTGTTGTACTTTTCAGTAC	101
QY	103	rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProTh	123
DB	102	TTTCGTTCTGTTGGCTCTTTCTGCTTCGGTGTGGCAGGTTATTTCATCAGTTGC	161

QY	123	rGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySe	143
DB	162	TCAAACAGGGATCATAAATGATTAGTCTCTCTGTTGCACATACTCCATGTTGGTTC	221
QY	143	rLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyri1	163
DB	222	ATATCATGACTTTTGGAGGAATGATTGGTGCATCTTCAGCGGGAAGTTCAGATCTCAT	281
QY	163	eGlyArgLysGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAlaI1	183
DB	282	GGGTGAAAAGGAGCTATGTTGGTTCCTCAAAATTTCTGCATCTTCGTTGGTTCAGT	341
QY	183	eSerPheAlaLysAspAlaSerPheLeuTyMetGlyArgLeuLeuGluGlyPheGlyVa	203
DB	342	AGCATTAGCAAAAGACTCCATGTGGCTTGATATTGGAAGACTATCCACAGGATTTCAGT	401
QY	203	lGlyIleIleSerTyThrValProValTyriIleAlaGluIleSerProGlnAsnMetAr	223
DB	402	TGTTTTATTAAAGCTATGTATACCATCCAGTTTACATTCAGAAATTAACCAACATGTTTCG	461
QY	223	gGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyriLe	243
DB	462	AGGAGCGTTTGTATTGCTAATCAGCTGATGCAGAGTTGTGGATTGCTTTATTCTACGT	521
QY	243	uLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMe	263
DB	522	CATTGGAAATTTTGTTCATTGGCGTAACCTGGCGCTTAATCGGCTCATCTCCATGTCGTT	581
QY	263	tLeuIleProGlyLeuPheIleProGluSerProArgTrpLeuAlaLysMetAsnLe	283
DB	582	GCAAGTTGTGACTTTTCTTTTATCCAGAGTCCCTAGACTACTGGGAAAATGGGACA	641
QY	283	uThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrTh	303
DB	642	TGAAAAGAAATGTAGAGCTTCATTGCAAACTCTTCGGGAGAGATGATGCAGATATCTCTGA	701
QY	303	rGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPh	323
DB	702	AGAAGCAACACTATCAAAAGAACCATGATCTTGTGTGATGAAGGACCAAAATCGCGGT	761
QY	323	eGlnGluLeuAsnGlnLysTyArgThrProLeuLeuLeuGlyIleGlyLeuLeuVa	343
DB	762	TATGGATTGTTTTCAGAGAAGATATCTCCATCTGTTGTTATTTGGTGGGACTAATGCT	821
QY	343	lLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyAlaSerSerIlePheLysAl	363
DB	822	TCTACAACACTCTCTGGAAGCTCAGGACTTATGTACTATGTCGTTAGCGTATTTCATAA	881
QY	363	aAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAl	383
DB	882	AGGAGGGTTTCCAAGCAGC---ATTGGCTCAATGATTTCTTGCAGTGCATCATGATACCAA	938
QY	383	aThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuIleIleSe	403
DB	939	AGCTCTATTGGGTCTGATTTGGTTGAGAAAATGGGACCAAGACCACTTCTATTGGCCTC	998
QY	403	rThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAs	423
DB	999	TACCGGTGGAATGTGCTTTTTCAGCTTGTCTCTCAGTTTTCCTTCTGCTTTCGG-----	1053
QY	423	nIleSerGlnAspSerAsnSerTyTyriIleLeu-----ThrMetIleSerLe	439
DB	1054	-----TCATATGGCATGCTTGATGAGTCACTCCGATTTTTCATG	1094
QY	439	uValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLe	459
DB	1095	TATCGGTGATGGGGTTTCATCTCTTCATTTGCCGTAGGCATGGGAGGCTTACCATGGAT	1154
QY	459	uMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLe	479
DB	1155	CATCATGTCAGATATTCCCAATGAATGTTAAAGTTTCTGCTGGGACTCTGTTACCTT	1214

Fri Apr 15 06:53:45 2005

```
Qy 479 uAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSe 499
Db 1215 AGCCAACTGGTCCTTGGTGGATGGTGGCTTCGCCCTACAACTTCATGCTAGAGTGGAA 1274
Qy 499 rValGlyGlyThrPheLeuSerTyzMetValValSerAlaPheThrIleValPheValVa 519
Db 1275 CGCATCAGGACGTTCTTGAATCTTCTTACTTCTTACTTCTTACTTCTTACTTCTTACTT 1334
Qy 519 lLeuTrpValProGluThrLysGly 527
Db 1335 TGCATGGTACCCGAACTAAAGGA 1359
```

Search completed: April 13, 2005, 15:31:04  
Job time : 668.423 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:36:34 ; Search time 223.162 Seconds  
(without alignments)  
4054.730 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779

Sequence: 1 PSSSSFRPAGKKKKKQKQ.....TIFVSLSIORQLWLPECLLS 553

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool.p/US10051909/runat\_13042005\_074038\_14080/app.query.fasta\_1.1678  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0 1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN 1 141 @runat\_13042005\_074038\_14080 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	578	20.8	5228	4	US-09-919-039-216
2	560.5	20.2	5227	4	Sequence 216, App
3	557.5	20.1	2089	3	Sequence 79, Appl
4	557	20.0	2017	3	Sequence 25, Appl
5	551.5	19.8	3915	4	Sequence 21, Appl
6	542.5	19.5	1853	3	Sequence 1104, App
7	528	19.0	2592	3	Sequence 23, Appl
8	527.5	19.0	1914	3	Sequence 8, Appl
9	524	18.9	2592	4	Sequence 19, Appl
10	523.5	18.8	2856	3	Sequence 8, Appl
11	523.5	18.8	2856	4	Sequence 135, App
12	523.5	18.8	2856	4	Sequence 135, App

13	523.5	18.8	2856	4	US-09-606-421B-135	Sequence 135, App
14	523.5	18.8	2856	4	US-09-221-107-135	Sequence 135, App
15	523.5	18.8	2856	4	US-09-466-396A-135	Sequence 135, App
16	523.5	18.8	2856	4	US-09-476-496A-135	Sequence 135, App
17	523.5	18.8	2856	4	US-09-630-940B-135	Sequence 135, App
18	523.5	18.8	2856	4	US-09-285-479-135	Sequence 135, App
19	522	18.8	1872	3	US-09-291-922-27	Sequence 27, Appl
20	519	18.7	3000	2	US-08-528-692-9	Sequence 9, Appl
21	519	18.7	3000	3	US-09-339-972-9	Sequence 9, Appl
22	514.5	18.5	1431	4	US-09-489-039A-4762	Sequence 4731, App
23	497	17.9	1545	4	US-09-489-039A-4761	Sequence 17, Appl
24	470.5	16.9	1752	4	US-09-679-686B-17	Sequence 11, Appl
25	469	16.9	1776	4	US-09-679-686B-11	Sequence 15, Appl
26	455	16.4	1960	4	US-09-679-686B-15	Sequence 1, Appl
27	451	16.2	1675	4	US-09-679-686B-1	Sequence 1, Appl
28	450	16.2	1704	4	US-09-614-221A-96	Sequence 96, Appl
29	447.5	16.1	1668	4	US-09-614-221A-420	Sequence 420, App
30	446.5	16.1	1704	4	US-09-614-221A-534	Sequence 534, App
31	446	16.0	1506	4	US-09-489-039A-4560	Sequence 4560, App
32	442	15.9	1659	4	US-09-248-796A-6610	Sequence 6610, App
33	439.5	15.8	987	4	US-09-710-279-1941	Sequence 1941, App
34	439.5	15.8	3065	4	US-09-710-279-3697	Sequence 3697, App
35	439.5	15.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
36	435	15.7	740	4	US-09-949-016-3739	Sequence 3739, App
37	435	15.7	1566	4	US-09-489-039A-2378	Sequence 2378, App
38	421.5	15.2	1626	4	US-09-614-221A-521	Sequence 521, App
39	420	15.1	1515	4	US-09-248-796A-6607	Sequence 6607, App
40	418	15.0	1714	4	US-09-774-528-437	Sequence 437, App
41	417	15.0	1943	4	US-09-774-528-168	Sequence 168, App
42	416.5	15.0	1689	4	US-10-162-012-45	Sequence 45, Appl
43	416.5	15.0	4167	4	US-09-949-016-4759	Sequence 4759, App
44	416.5	15.0	4385	4	US-10-162-012-43	Sequence 43, Appl
45	413.5	14.9	1521	4	US-09-248-796A-5972	Sequence 5972, App

#### ALIGNMENTS

RESULT 1

US-09-919-039-216

; Sequence 216, Application US/09919039

; Patent No. 6727066

; GENERAL INFORMATION:

; APPLICANT: Kaser, Matthew R.

; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

; FILE REFERENCE: PA-0035 US

; CURRENT APPLICATION NUMBER: US/09/919,039

; PRIOR FILING DATE: 2002-09-09

; PRIOR APPLICATION NUMBER: 60/222,113

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 401

; SOFTWARE: PERL Program

; SEQ ID NO 216

; LENGTH: 5228

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6727066 410785.1

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 4929, 4935, 4940, 4945, 4974, 4993

; OTHER INFORMATION: a, t, c, g, or other

US-09-919-039-216

Alignment Scores:

Pred. No.: 2.13e-52 Length: 5228

Score: 578.00 Matches: 152

Percent Similarity: 49.22% Conservative: 101

Best Local Similarity: 29.57% Mismatches: 167

Query Match: 20.80% Indels: 94

DB: 4 Gaps: 15

US-10-051-909-36 (1-553) x US-09-919-039-216 (1-5228)



105 IleValAla---LeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProThr 123  
 81 ATCATGCTGCTGCTGGTTCCTTCCAGATTGGATATGACATTGGTGTGATCAATGCACCT 140  
 124 GlnAspAlaMetValArgAspLeu----- 131  
 141 CAACAGGTAAATATATCTCACTATAGACATGTTTGGGTGTTCCACTGGATGACCGAATA 200  
 131 ----- 131  
 201 GCTATCAACAACATGTTATCAACAGTACAGATGAACCTGCCACAATCTCATACTCAATG 260  
 131 ----- 131  
 261 AACCCAAAACCAACCCCTGGGCTGAGGAAGAGACTGTGGCAGCTGCTCAATAATCAACC 320  
 132 -----AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGly 148  
 321 ATGCTCTGGTCCCTGCTGTCATCCAGCTTGCA-----GTTGGT 359  
 149 GlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGlySer 168  
 360 GGAATGACTGCATCATCTTCTGGTGGTGGCTGGGACACACTTGGAGAATCAAAAGCC 419  
 169 LeuMetIleAlaIleAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaIys--- 187  
 420 ATGTTAGTACAAACATCTGCTATAGTTGGAGCTCTCTTGATGGGTTTTCAAAATTG 479  
 188 -----AspAlaSerPheLeuTyrMetGlyArgLeuGluGlyPheGlyValGlyIle 205  
 480 GGACCATCTCATATATTAATGTTGGTGAAGAGCATATCAGGACTATATTTGGGCTA 539  
 206 IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225  
 540 ATTTCAGGCCCTGCTCTATGATATATCGGTGAATTTGCTCAACCGCTCTCAGGGGAGCA 599  
 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe----- 239  
 600 CTTGGCATTCTTTCAGCTGGCCATCGTCACCGGCATCTTATTAGTCAGATTATTTGGT 659  
 240 LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu 259  
 660 CTTGAATTTATCTGGCAATTATGATCTGTGGACATCTTGCTGGCCCTGCTGTGGTGTG 719  
 260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu--- 278  
 720 CGAGCCATCTCTCAGTCTCTGCTACTCTTTTCTGTCAGAAAGCCCGACATACCTTTAC 779  
 279 AlaIysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298  
 780 ATCAAGTTAGATGAGGAAGTCAAGCAAAACAAAGCTTGAAGAAGACTCAGAGGATATGAT 839  
 299 ThrAspIleThrGluValAsnAspIleIys-----ArgAlaValAlaSerSerSer 316  
 840 ---GATGTCAACCAAGATATAATTAATGAATGAGAAAGAAAGAGAAAGAGCATCGATGAG 896  
 317 LysArgThrThrIleSerPheGlnGluLeuAsnGlnLysIysTyrArgThrProLeuLeu 336  
 897 CAGAAAGTCTCTATA---ATTCAGCTCTTCCACCAATTCAGCTACCGACAGCTATTCTA 953  
 337 LeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyr 356  
 954 GTGGCACTGATGTCATGTGGCTCAGCAATTTTCCGGAATCAATGGCATTTTTTACTAC 1013  
 357 AlaSerSerIlePheIysAlaIleGlyValThrAsnSerAspLeuAlaThrCysSerLeu 376  
 1014 TCAACCAAGCATTTTTCAGACGGCTGGTATCAGCAAAACCTGTTATGCAACCATTTGGAGTT 1073  
 377 GlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArg 396  
 1074 GGGCTGTAAACATGGTTTCACTGCTGTCTGTATCTCTGTATCTCTGTGGAGAGGAGCGCGA 1133  
 397 ArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuLeuCysLeuLeuAlaValSerVal 416

1134 CGTTCTCTCTTCTTAATGGAATGATGGATGTTTGTGTGCGCATCTTCATGTCAGTG 1193  
 417 ValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMet 436  
 1194 GGACTTGTGCTGCTGAATAAGTTCTCTTG-GAC-----TGAGTTACT 1234  
 437 IleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIle 456  
 1235 GTGAGCATGATAGCCATCTCTCTTTGTGAGCTTCTTGAATTTGGCCAGCCGATC 1294  
 457 ProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIle 476  
 1295 CCTGTTTTCATGGTGGCTGAGTTTTTCAGTCAAGGACCACTCTCTGCTGCTGCTT 1354  
 477 AlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeu 496  
 1355 GCTGCATTTCAGCAATTTGGACCTGCAATTTCAATTTAGTACTCTGTGTTTCCAGTACAT 1414  
 497 ThrTrpSerValGlyGlyThrPheLeuSerTyrMet-----ValValSerAlaPhe 513  
 1415 GACTTCTGT-----GGACCTTATGTTTCTCTTCTGCTGGAGTGTCTCTGGCCCTTT 1468  
 514 ThrIleValPheValValLeuTrpValProGluThrLysGly 527  
 1469 ACCCTG---TTTACATTTTAAAGTTCCAGAAACCAAGGA 1507

RESULT 3  
 US-09-291-922-25  
 ; Sequence 25, Application US/09291922  
 ; Patent No. 6383776  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Steve  
 ; APPLICANT: Hitz, Bill  
 ; APPLICANT: Kinney, Tony  
 ; APPLICANT: Tingey, Scott  
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins  
 ; FILE REFERENCE: BB-1163  
 ; CURRENT APPLICATION NUMBER: US/09/291,922  
 ; CURRENT FILING DATE: 1999-04-14  
 ; EARLIER APPLICATION NUMBER: 60/083,044  
 ; EARLIER FILING DATE: April 24, 1998  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 25  
 ; LENGTH: 2089  
 ; TYPE: DNA  
 ; ORGANISM: Triticum aestivum  
 US-09-291-922-25

Alignment Scores:  
 Pred. No.: 7,78e-51 Length: 2089  
 Score: 557.50 Matches: 135  
 Percent Similarity: 49.79% Conservative: 97  
 Best Local Similarity: 28.97% Mismatches: 187  
 Query Match: 20.06% Indels: 47  
 DB: 3 Gaps: 7

US-10-051-909-36 (1-553) x US-09-291-922-25 (1-2089)

QY 99 AlaPheLeuCysThrThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
 DB 137 GCCTTCGCTGCGCCATCTCGCTCCATGACCTCCATCTCTCTGGCTACGACATCGCG 196  
 QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138  
 DB 197 GTGATGAGCGGAGCGTGGTGTATCTCCAGAGAGATCTGAAGATCAACGACACCCAGCTG 256  
 QY 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158  
 DB 257 GAGGTCTCTCATGGGCATCTCCACGCTGATCTGCTCATTTGGCTCTCTTCGGCGGGCGG 316  
 QY 159 MetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaIleProAsnIleIle 178

```

Db      317  AGTCCGATCGGCGCGGCTTCCACATCGTCTTCGCGCGCTCTTCTTCGGG 376
Qy      179  GlyThrPheLeuSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198
Db      377  GCGCGCTTCATCGGCTTCTCCGTCACATACGCGCATGCTCATGTTCGGCGCGTTCGTG 436
Qy      199  GluGlyPheGlyValGlyLeuLeuSerTyrThrValProValTyrIleAlaGluLeuSer 218
Db      437  GCGCGCATCGGCGTGGGTGAGCTCTCATGATCGCGCGCTGAGACACGCGGAGGTGTC 496
Qy      219  ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238
Db      497  CCGCGCTCGCGCTGGGTCTCACATCTCTCCGCGAGGTTCATCACTTCGCGATC 556
Qy      239  PheLeuAlaTyrLeu-----LeuGlyMetPheIleProTyrArg 251
Db      557  CTCCTCGGATATGCTCCAACTTCGCTCCCGCGCTCTCCCTCCCGCTCGGCTGCGC 616
Qy      252  LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuLeuProGlyLeuPheIle 271
Db      617  ATTATGCTCGGATAGCGCGGTGCGCTCGCTCGCTTCGCTTCATGGTTCGCGCATG 676
Qy      272  ProGlnSerProArgTyrPheLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db      677  CCGAGTCTCCCGGTGGCTGCTCATGAGGCGGCTCTCGCGGACGCAAGGTGTGCTT 736
Qy      292  GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db      737  GCCAAGACGTCGACACACGCGAGAGGCGCGGAGGCGCATCGCGGACATTAAAGCTGC 796
Qy      312  Val-----Ala 313
Db      797  GCGGCGATCCCTCTGGGCTCGACGCGGACGTGTCCCGTCCCGTCCCAAAAACAAGGAGC 856
Qy      314  SerSerSerLysArgThr-----ThrIleSerPheGlnGlu 325
Db      857  ACGAGGAGAGCGGCTTTGAGAGACCTCATCTGTACCGACCATAGCATCGGCCAC 916
Qy      326  LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGln 345
Db      917  ATC-----CTCATCGGGAATCGGCATCCACTCTCTTCCAG 952
Qy      346  AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaGly 365
Db      953  CAGTCTTCGGGATCGACCGCTGTGCTCTACAGCCCGCTAGTATTTTCAAGAGCGCGGC 1012
Qy      366  ValThrAsnSerAspLeu-----AlaThrCysSerLeuGlyAlaIleGlnValLeu 382
Db      1013  ATCAGCGGCGACAGCGCTCTCGCGGACACACCGTGGCGGTTCGGGCGACCAATACGGTC 1072
Qy      383  AlaThrGlyValThrThrTyrLeuLeuAspArgAlaGlyArgArgIleLeuLeuIle 402
Db      1073  TTCATCTGGTGCCACCTCTCTCGACCGCATCCGCGGCGCGCTGGTGTGCTGACC 1132
Qy      403  SerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAsp 422
Db      1133  ACACGCGCGCATGCTCTCTCTAGTGGCTTCGCGCGGCTTCACCGGTATCAGC 1192
Qy      423  AsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIle 442
Db      1193  CGCCACCGCGAGAGAGATCCTTCGCGGCGCATC-----GTCCTGTGCATCTTCTGCATC 1246
Qy      443  ValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMetSer 462
Db      1247  ATGGCTACGTGGCTCTCTCTCCATCGGCGCTCGGCGGCGGCGGCGGCGGCGGCGG 1306
Qy      463  GluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTyr 482
Db      1307  GAGATCTTCCGCTCGACGTGGCGGCTGGGCTGCTCCCTCGGCGGCGGCGGCGGCGG 1366
Qy      483  LeuThrSerPheAlaIleThrMetThr-----ThrAsnLeuMetLeuThrTyrSerValGly 501

```

```

Db      1367  CTGACCGCGCGGTGATCTCCATGACCTTCATTTTCGCTGCCAAGGCCATGACCATCGGC 1426
Qy      502  GlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValLeuTyr 521
Db      1427  GCGCGCTTCTTCCTTCCTTCGCGCGCATCGCTCATTCGATGGGTGTTCTTCTTCGCGCTAC 1486
Qy      522  ValProGluThrLysGly 527
Db      1487  CTGCGGAGACCCGCGC 1504

RESULT 4
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-21

Alignment Scores:
Pred. No.:      8,31e-51      Length:      2017
Score:          557.00      Matches:      148
Percent Similarity: 49.01%      Conservative: 99
Best Local Similarity: 29.37%      Mismatches: 211
Query Match:      20.04%      Indels:      46
DB:                3          Gaps:          9

US-10-051-909-36 (1-553) x US-09-291-922-21 (1-2017)

```

```

Qy      80  SerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer--- 98
Db      75  TCCTCCTCTGCACCAACCGGAGATGGCTTCGCGCGCATCTCGCTCCATGACCTCATC 134
Qy      99  -----AlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db      135  AAGAAGGGCAACGTCGCGTTCGCTTCGCTGCCGCCATCTCGCTCCATGACCTCATC 194
Qy      112  GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db      195  CTCCTCGGCTACGATATCGGGGTGATGAGCGGGCGTTCGCTGTACATCAAGAAGGACTTC 254
Qy      132  AsnLeuSerIleSerGlyPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db      255  AACATCATGTGACGGGAAGGTGGAGTTCATGGGCATCTGAACCTCTACTCGCTCATC 314
Qy      152  GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db      315  GGTCTTCGCGCGCGGCGGACGTCGATCGATCGCGCGGCGGTACACCATCGTGTTC 374
Qy      172  AlaIleProAsnIleIleGlyTyrPheAlaIleSerPheAlaLysAspAlaSerPhe 191
Db      375  GCGCGGTCAATTTCCTCGCGGGGGGTTCCTCATGGGTTCGCGCTCAACTACGCATG 434
Qy      192  LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db      435  CTCATGTTCGCGCGCTTCGTGGCGCGCATCGGCTGGGTTCACGCGCTCATGATCGCGCG 494
Qy      212  ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231

```



Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy Db Qy

495 GTGTACACCGCCGAGGTGTGCCGGCGTTCGGCGGTGCCTCCTATGACTGTTCTTCCGCGAG 554  
232 LeuSerValThrPheGlyIlePheLeuAlaTyrlau-----Leu 244  
:: :: :: :: :: :: :: :: :: :: :: ::  
555 GTGTTTCATCAACTTCGGCATCTGCTCGGTACGTCTCGAACTATGCTTTCTCCGCGTTG 614  
245 GlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeu 264  
:: :: :: :: :: :: :: :: :: :: :: ::  
615 CCGCTGAACCTCGGGTGCGGCATCATCTCGGCATCGCGCGCGCGTCCGTGCTGCTGCTC 674  
265 IleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThr 284  
:: :: :: :: :: :: :: :: :: :: :: ::  
675 GCGCTCATGTGTCTCGCATGCCGAGTCCCGCGGTGGTGGTGTATGAAGGAGCGGCTC 734  
285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGlu 304  
:: :: :: :: :: :: :: :: :: :: :: ::  
735 GCGGACGCCAAGGTGGTGTCTGGAGAAGACCTCCGACACGGCGGAGGAGCGCGGAGCGC 794  
305 ValAsnAspIleLysArgAla-----ValAlaSer 314  
:: :: :: :: :: :: :: :: :: :: :: ::  
795 CTGGCCGACATCAAGGCGCGCGCGCATCCCTGAGGAGCTCGACGCGCAGTGGTGTCACC 854  
315 SerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrlauPro 334  
:: :: :: :: :: :: :: :: :: :: :: ::  
855 GTCCCCAAGAGAGGAGCGGGAACAGACGAGCGGTGTGGAGGAGCTCATCTGTCTCCCG 914  
335 -----LeuLeuGlyIleGlyLeuValLeuGlnAsnLeu 347  
915 ACCCGGCCATCGGCGCATCTGCTGTCGGGATCGCATCCACTTCTTCCAGCATGCG 974  
348 SerGlyIleAsnGlyValLeuPheTyrlauSerSerIlePheLysAlaAlaGlyValThr 367  
:: :: :: :: :: :: :: :: :: :: :: ::  
975 TTGGGCATTCACTCGGTGCTTCTACAGCCCTCTGTTTCAAGAGCCCGGATTAAACG 1034  
368 AsnSerAspLeu-----AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThr 384  
1035 AAGGACAAACACTTCTTGGGCACCACTTGGCGGTTCGTGTCACCAAGAGCTTTTCATC 1094  
385 GlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuIleIleSerThr 404  
:: :: :: :: :: :: :: :: :: :: :: ::  
1095 TTGTTGGCGACTTCTTCATCGACGGGTGCGGCGCGCGCTGTTGCTGGGCGAGCACG 1154  
405 SerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIle 424  
:: :: :: :: :: :: :: :: :: :: :: ::  
1155 GGCGGATAATCTCTCTCCCTCATCGGCTCGGCGCGGCTCACCCTC-----GTC 1205  
425 SerGlnAspSerAsnSerTyrlauLeuThrMet---IleSerLeuValGlyIleVal 443  
1206 GCCGACACCCCGACCGCAGATACCTTGGGCGCATCGGCTTAAGCATCGCTCCACCCTC 1265  
444 SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlu 463  
1266 GCCTACGTGCGCTTCTTCTCATCGGCTTGGCCCATCATCGTGGGTGCACAGCTCGAG 1325  
464 IleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeu 483  
1326 ATCTTCCGCTCCAGGTGCGGCGCTGGGCTGCTCGCTCGGCGTCCGCGCAACCGCGTC 1385  
484 ThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThrTrpSerValGlyGly 502  
1386 ACCAGCGGCGTCACTCTCCATGACCTCTGCTGCTGTCACCAAGGCGCATCAGTCGGCGGC 1445  
503 ThrPheLeuSerTyrlauMetValValSerAlaPheThrIleValPheValValLeuTrpVal 522  
1446 AGCTTCTTCTTACTCGGCGCATCGGCGGCTCGCTGGGTGTTTCTTACACCTACTCCTC 1505  
523 ProGluThrLysGly\*\*\*AsnSerArg-----GlyAspThrIlePhe 536  
1506 CCGGAGACCCGCGCGGAGACGCTCGAGGAGATGAGCAAGCTGTTTCGGCGCACGCGCGCC 1565  
537 ValSerLeuSer 540  
1566 GCCTCGGAATCA 1577

## RESULT 5

US-09-023-655-1104  
 ? Sequence 1104, Application US/09023655  
 ? Patent No. 6607879  
 ? GENERAL INFORMATION:  
 ? APPLICANT: Cocks, Benjamin G.  
 ? APPLICANT: Susan G. Stuart  
 ? APPLICANT: Jeffrey J. Seilhamer  
 ? TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 ? TITLE OF INVENTION: EXPRESSION  
 ? NUMBER OF SEQUENCES: 1508  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ? STREET: 3174 PORTER DRIVE  
 ? CITY: PALO ALTO  
 ? STATE: CALIFORNIA  
 ? COUNTRY: USA  
 ? ZIP: 94304  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/09/023,655  
 ? FILING DATE: HEREWITH  
 ? CLASSIFICATION:  
 ? PRIOR APPLICATION DATA:  
 ? APPLICATION NUMBER:  
 ? FILING DATE:  
 ? CLASSIFICATION:  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Zeller, Karen J.  
 ? REGISTRATION NUMBER: 37,071  
 ? REFERENCE/DOCKET NUMBER: PA-0001 US  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: (650) 855-0555  
 ? TELEFAX: (650) 845-4166  
 ? INFORMATION FOR SEQ ID NO: 1104:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 3915 base pairs  
 ? TYPE: nucleic acid  
 ? STRANDEDNESS: single  
 ? TOPOLOGY: linear  
 ? IMMEDIATE SOURCE:  
 ? LIBRARY: GENBANK  
 ? CLONE: g183684  
 ? US-09-023-655-1104

Alignment Scores:	
Pred. No.:	1.08e-49
Score:	551.50
Percent Similarity:	48.95%
Best Local Similarity:	30.59%
Query Match:	19.95%
DB:	4
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:
	3915
	145
	87
	180
	62
	13

US-10-051-909-36 (1-553) x US-09-023-655-1104 (1-3915)

Qy	94	GlusSerHisValSerAlaPheLeuCys--ThrLeuIleValAlaLeuGlyProIleGlnPh	113
Db	254	GAAGGTCAACCCAGCTCTGATATTGGCCATCACAGTTGCTACAATCGGCTCTTTCCAA	313
		:    :    :    :    :    :    :	
Qy	113	eGlyPheThrSerGlyPhe-----	119
		:    :    :    :    :    :    :	
Db	314	TGGCTACAACACTCGGGTCATCAATGCTCCTGAGAAGATCATAAAGGAATTTATCAATAA	373
		-----SerSerProThrGlnAspAlaMetValAtgAspLeu---As	
Qy	120	-----	
		:    :    :    :    :    :    :	
Db	374	AACTTTGACGGACACAGGGAATGCCCAACCCTCTGAGGTGCTGCTCACAGCTCTCTCTGGGC	433
		-----	
Qy	132	nLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAanValGlyGlyMetValGI	152

```

Db 434 CTTGCTGTGGCCATATTTTC-----TTTTTTTTTTTTTTTGGG 472
Qy 152 YAlaileAlaserGlylnMetAlaGluTyrileGlyArglySglsrLeuMetileAl 172
Db 473 CTCTTTTTCGTCGGACTCTCTGTCACCGCTTTTGGCAGGCGCAATTCATGCTGATTGT 532
Qy 172 aAlaileProAsnileileGlyTrpLeuAlaileSerPhe-----AlaLysAspAl 189
Db 533 CAACCTGTGTGGCTGCTCAGTGGTGGCTGCTTATGGGACTGTGTAAGTAGCTAAGTCGGT 592
Qy 189 aSerPheLeuTyrMetGlyArgLeuGluGlyPheGlyValGlylleileSerTyrTh 209
Db 593 TGAATGCTGATCTCGTGGTGGCTTGTGTTATGGCTCTTCGCGGACCTGCACAGGTTT 652
Qy 209 rValProValTyrileAlaGluileSerProGlnAsnMetArgGlyAlaLeuGlySerVa 229
Db 653 TGTGCCCATGTATCTTGGAGAGATCTCGCTACTCCCTGGCGGTGCTTTGGCATCT 712
Qy 229 lAsnGlnLeuSerValThrPheGlyllePheLeuAlaTyrLeuGlyMet----PheIl 248
Db 713 CAACCACTGGGATCGTTGTTGGAATTCGTGGTGGCCAGATCTTGTGCTGGAAATTCAT 772
Qy 248 e-----ProTrpArgLeuLeuAlaValileGlyAlaLeuProCysThrMe 263
Db 773 CTTGGGCTGAGAGCTATGCGCCGCTACTGCGGTTTATACCATCCTCTCTGCTATCCT 832
Qy 263 tLeuileProGlyLeuPhePheileProGluSerProArgTrpLeuAlaLysMetAsnLe 283
Db 833 ACAAGTCGAGCCCTCCATCTTCCATTTGCTGAAAGTCCAGATTTTGTGCTC---ATTAACAG 889
Qy 283 uThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe-----GluThrAspIl 301
Db 890 AAAAGAAGAG---GAGAATGCTTAAGCAGATCTCCAGCGGTTGTGGGCGCCAGGATGT 946
Qy 301 eThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrIl 321
Db 947 ATCCCAAGACATCCAGGAGATGAAGATGAGATGCAAGATGTCACAGAAAGCAAGT 1006
Qy 321 eSerPheGlnGluLeu---AsnGlnLysLysTyrArgThrProLeuLeuGlylleGl 340
Db 1007 CACCGTGCTAGAGCTCTTTAGAGTGTCAGCTACCGACGCCATCAATTCATTCATGT 1066
Qy 340 yLeuLeuValLeuGlnAsnLeuSerGlylleAsnGlyValLeuPheTyrAlaSerSerIl 360
Db 1067 GCTCCAGCTCTCTCAGCAGCTCTCTGGGATCAATGCTGTCTTATTACTCAACAGGAAT 1126
Qy 360 ePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaileGl 380
Db 1127 CTTCAAGATGAGGTGTTCAAGAGCCCATCTATGCCACCATCGCGCGGCTGTGGTTAA 1186
Qy 380 nValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgGileLeuLe 400
Db 1187 TACTATCTCACTGTAGTTTCTCTATTCTGTTGGTAAAGGCGAGGAAGAGGACTCTGCA 1246
Qy 400 uileileSerThrSerglyMetThrLeuCysLeuLeuAlaValSerValValPhePheLe 420
Db 1247 TATGATAGCGCTTGAGGGATGGCTTTTGTTCACGCTCATGACTGTTCTTTGTATT 1306
Qy 420 uLysAspAsnileSerGlnAspSerAsnSerTyrTyrileLeuThrMetileSerLeuVa 440
Db 1307 AAAGATTAAC-----TATAATGGGATGAGCTTTGCTGTGATTGG 1345
Qy 440 lGlylleValSerPheValileThrPheSerPheGlyMetGlyAlaileProTrpLeuMe 460
Db 1346 GGCTATCTTGTGCTTTGTAGGCTTCTTTGAAATGGACCGAGGCCCATTCCTCGGTTTAT 1405
Qy 460 tMetSerGluLeuLeuProValSerileLysSerLeuGlyGlySerileAlaThrLeuAl 480
Db 1406 TGTGGCCGAATCTTTCAGCAGGCGCCCGCCAGCTGCGATGCGAGCGCGGCTGCTC 1465
Qy 480 aAsnTrpLeuThrSerPheAlaileThrMetThrThrAsnLeuMetLeuThrTrpSerVa 500
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

```

```

Db 1466 CAACCTGGACCTCCAACTTCCTAGTCGGA-----TTGCTCTTCCCTC 1507
Qy 500 lGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThr----- 514
Db 1508 CGCTGCTCACTATTATTAGGAGCTAGCTTTTATTATCTTCACGGCTTCTCTATTACCTT 1567
Qy 515 -ileValPheValValLeuTrpValProGluThrLysGly 527
Db 1568 CTTGGCTTTTACCTTCTTCAAAGTCCCTGAGACCCGTTGC 1607

RESULT 6
US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23

Alignment Scores:
Pred No.: 2,86e-49 Length: 1853
Score: 542.50 Matches: 137
Percent Similarity: 50.64% Conservative: 102
Best Local Similarity: 29.03% Mismatches: 183
Query Match: 19.52% Indels: 51
DB: 3 Gaps: 9

US-10-051-909-36 (1-553) x US-09-291-922-23 (1-1853)
Qy 93 ArgGluSerHisValSerAlaPheLeuCysThrLeuileValAlaLeuGlyProIleGln 112
Db 239 CGCAAAAGGAACAAGATGCTTTTGTGCTATGCTGCGCTCCCATGACTTCCATCTTG 298
Qy 113 PheGlyPheThrSerglyPheSerProThrGlnAspAlaMetValArgAspLeuAsn 132
Db 299 CTTGGTTATGATATTGGAGTGATGATGGAGCAGCCATATACATAAAAGGACCTGAAA 358
Qy 133 LeuSerileSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
Db 359 GTCTCGGACGAGCAATCGAGATCTGCTCGGAATCATCACTACTCTCTGATAGGC 418
Qy 153 AlaileAlaSerGlnMetAlaGluTyrileGlyArglySglsrLeuMetileAla 172
Db 419 TCATGCTCTCGCGGCAACCTCCGACTCGATAGTCCCGGTACAGATTGTTTTCGCC 478
Qy 173 AlaileProAsnileileGlyTrpLeuAlaileSerPheAlaLysAspAlaSerPheLeu 192
Db 479 GGCACCACTCTTGTGCGGAGCATTCTCATGGGTCTCTCCCAATTTATCTTCTTC 538
Qy 193 TyrMetGlyArgLeuLeuGluGlyPheGlyValGlylleileSerTyrThrValProVal 212
Db 539 ATGTTTTCGCGTTTCGTCGCTGGCATTTGGCATGCGGCTACGCGCTCATGATAGCCCCCGTC 598
Qy 213 TyrileAlaGluileSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu 232
Db 599 TACACCGCCGAGGCTCTCCCGGCTCTCTGCTGGGCTTCTCACTTCTCTTCCCTGAGGTA 658
Qy 233 SerValThrPheGlyllePheLeuAlaTyrLeu-----LeuGly 245
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

```

```

Db 659 TTATTATTCGAGGATTAATTTGGATACATATCAAACTATGCATTTTCGAGCTGACA 718
Qy 246 MetPheLeuProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuLe 265
Db 719 CTAAGGTGGGATGGGCAATGATGCTTGGAGTTGGTGGCAATACCTTCGGTACTCTTAACA 778
Qy 266 ProGlyLeuPhePheLeuProGlySerProArgTrpLeuAlaLys----- 280
Db 779 GTAGGAGTGTGGCGATCGCGAGTCCCAAGGTGGCTTGTGATGAGGGTCTGTTGGGA 838
Qy 281 -----MetAenLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295
Db 839 GAGCAAGAAAGTGTCTTAACAACTCAGACGACGAGCAAGAGCCCACTA----- 892
Qy 296 GlyPheGluThrAspLeuThrGluValAsnAspIleLysArgAlaVal-AlaSerSe 315
Db 893 -----AGGCTAGCGGAAATCAACAAGCCGAGGATCCCC 928
Qy 315 rSerLysArgThrIleSerPheGlnGluLeuAenGln----- 328
Db 929 GAGAGTTGCACGACGCGTTCAGGTAAATAAACAAGCAACGCGTGAAGTGTATGG 988
Qy 329 -----LysLysTyArgThrPro-----LeuLeuLeuGlyIleGl 340
Db 989 AAAGAGCTCTTCCTC-TATCCACGCGCGCAATTGTCACATCGTAATCGTCCCTTGG 1047
Qy 340 yLeuLeuValLeuGlnAsnLeuSerGlyIleAenGlyValLeuPheTyTrpAlaSerSeril 360
Db 1048 TATTACTCTTCTTCAACAAGCGTGGCGGTAGACGCGTCTGTTGTACAGCCCAAGAT 1107
Qy 360 ePheLysAlaAlaGlyValThrAenSerAsp-----LeuAlaThrCysSerLeuGl 377
Db 1108 CTTTCAAGAGCGTGGGATTAACAACACGACGCAATAAGCTTCTTGCACACGCGTGG 1167
Qy 377 yAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgAr 397
Db 1168 ATTCTGTAAAGCGTGTTCATCTTGGCGGTACGTTTACGTTGGACGCGTGGTGTGTCG 1227
Qy 397 gIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVa 417
Db 1228 TCGTGTGTTATTGCTAGTGTGGCGCATGGTCTCTCGCTTCTCACGCTTGGATCAG 1287
Qy 417 lPhePheLysAspAsnIleSerGlnAspSerAenSerTyTrpIleLeuThrMetIl 437
Db 1288 CTTCACTGTATTGATCAT---TCGAGAGAGAAATTAATGTGGCGGTGGATCGAGCAT 1344
Qy 437 eSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIlePr 457
Db 1345 AGCATGGTG-----TTGGCTTACGTGGCCAGTCTCATCGTGGGTCCCATCAC 1398
Qy 457 ofTrpLeuMetMetSerGluLeuLeuProValSerIleLysSerLeuGlyGlySerIleAl 477
Db 1399 GTGGGTCTATAGTCTGAGATCTCCGTTGAGGCTGCGGCGCARGGTGCGCGCGGG 1458
Qy 477 aThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr---ThrAenLeuMetLe 496
Db 1459 AGTTGGGTGGAATAGGACCACTAGCGGGTGTCTCAATGACTTTTCTGTCCTCACTAG 1518
Qy 496 uThrTrpSerValGlyGlyThrPheLeuSerTyMetValValSerAlaPheThrIleVa 516
Db 1519 AGCCATCACTATTTGGTGGACTTCTTCCTTAITTTGGGATTTGCTACTGTTGGTGGAT 1578
Qy 516 lPheValValLeuTrpValProGluThrLysGly 527
Db 1579 ATTCTTTTACACGCTCTTGCTCGAGACCGGGGA 1612

```

RESULT 7

```

US-09-591-025-8
; Sequence 8, Application US/09591025
; Patent No. 6303373
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; LODISH: Lodish, Harvey F.

```

```

; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; TITLE OF INVENTION: Targeting of GLUT4
; FILE REFERENCE: 0399.1210-004
; CURRENT APPLICATION NUMBER: US/09/591,025
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
US-09-591-025-8

```

```

Alignment Scores:
Pred. No.: 2,06e-47 Length: 2592
Score: 528.00 Matches: 131
Percent Similarity: 52.58% Conservative: 83
Best Local Similarity: 32.19% Mismatches: 165
Query Match: 19.00% Indels: 28
DB: 3 Gaps: 7

US-10-051-909-36 (1-553) x US-09-591-025-8 (1-2592)
Qy 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAenValGlyMetValGly 152
Db 562 CTCCTCGTGGCCATCTTTTCC-----GTGGCGGCATGATTTCC 600
Qy 153 AlaIleAlaSerGlyGlnMetAlaGluTyIleGlyArgGlySerLeuMetIleAla 172
Db 601 TCCTCTCTCATGTTGATCATCTCTCAGTGGCTTGAAGGAAAGGCGCATGCTGGTCAAC 660
Qy 173 AlaIleProAenIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSer----- 190
Db 661 AATGCTCTGGCGTGTGGGGGCGACGCTCATGGGCTGGCCCAACGCTGCTGCTCTAT 720
Qy 191 ---PheLeuTyMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyThr 209
Db 721 GAATGCTCATCTCTGGACGATTCCTCATTTGGGCGCTACTCAGGGCTGACATCAGGCTG 780
Qy 210 ValProValTyIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
Db 781 GTGCCCATGTACGTGGGGGAGATTGCTCCCACTCACCTGCGGGCGCGCTTGGGACGCTC 840
Qy 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTy-----Leu 243
Db 841 AACCAACTGGCCATTGTTATCGGCATTCTGATCGCCCGAGGTGCTGGGCTTGGAGTCCCTC 900
Qy 244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 901 CTGGGACCTGCCACGCTGTGGCCACTGCTCTGGGCGCTCACAGTGTACTTGCCTCCTG 960
Qy 264 LeuIleProGlyLeuPhePheIleProGlySerProArgTrpLeu---AlaLysMetAen 282
Db 961 CAGCTGCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy 283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAen 300
Db 1021 CTGAGGGGCTGCCCAAGAGATCTTGAAGCGCTGCACAGGCTGGGCGCGATGTTTCTGGA 1080
Qy 301 IleThrThrGluValAenAspIleLysArgAlaValAlaSerSerSerLysArgThrThr 320
Db 1081 GTCTGCTGAGCTGAAGATGAGAGCGGAAGCTG-----GAGCGTGAAGCGGCGCACTG 1134
Qy 321 IleSerPheGlnGluLeuAenGlnLysTyArgThrProLeuLeuLeuGlyIleGly 340
Db 1135 TCCCTGCTCCAGCTCTTGGCGACCGCTACCCACCGCGCGCTGATCATTTGGGCTGCTG 1194

```

341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle 360  
1195 CTGACGCTGAGCAGCAGCTCTCTGGCAATCAATGCTGTTTCTATTATTCACACGATC 1254  
361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380  
1255 TTCCGACAGCAGGAGGTAGCCAGCCTGCTATGACCACATAGAGCTGGTGGTCCAC 1314  
381 ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu 400  
1315 ACAGTCTTCCACTTGGTCTGCTGCTGTTGTGGAGCGGGCGGCGGACGCTCCAT 1374  
401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu 420  
1375 CTCTGGGCTGGCGGCATGTGTGCTGTGCCATCTCTGATGACTGTGGCTCTGCTCTG 1434  
421 LysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVal 440  
1435 CTGAGCGAGTCCA-----GCCATGAGCTACGTTCTCCATTGTG 1473  
441 GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMet 460  
1474 GCCATCTTGGCTTCTGTCGCAATTTTGTGATTTGGCTGGCCCTGCTTGGTTCATC 1533  
461 MetSerGluLeuLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAla 480  
1534 GTGGCCGAGCTTTCAGCCAGGAGACCCCGCCGCGGACCATGCTGTGCTGTTCTCC 1593  
481 AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal 500  
1594 AACTGGACGAGCAACTTCATCATTTGGCATGGGTTCCAGATGTTGCGGAGCTATGGG 1653  
501 GlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeu 520  
1654 CCCTACGCTCTCTCTTCTATTATGGGCTCTCTGCTGGGCTTCTTCATCTTCACTTCTTA 1713  
521 TrpValProGluThrLysGly 527  
1714 AGAGTACCTGAACTCGAGGC 1734

RESULT 8  
US-09-291-922-19 Application US/09291922  
; Sequence 19  
; Patent No. 638376  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1998-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 19  
; LENGTH: 1914  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-291-922-19

Alignment Scores:  
Pred. No.: 1,376-47 Length: 1914  
Score: 527.50 Matches: 132  
Percent Similarity: 48.72% Conservative: 97  
Best Local Similarity: 28.09% Mismatches: 186  
Query Match: 18.98% Indels: 55  
DB: Gaps: 7

US-10-051-909-36 (1-553) x US-09-291-922-19 (1-1914)

QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
Db 116 GCCTCATATGTGCATCTCTGGCTCCATGGCTCTGTCTCATCTTGTCAATGACATTTGGG 175  
QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138  
Db 176 GTGATGATGTGAGCGGCCCATGTACATCAAGAAGACCTGAATATCACGGACGTGACGCTG 235  
QY 139 SerAlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGln 158  
Db 236 GAGATCTGTATCGGGATCCTCAGTCTCTACTCGCTGTGTGGATCCTCTCGCGCGCGG 295  
QY 159 MetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIle 178  
Db 296 AGCTTCGACAGATCGGGCGCGCTTACCGCTGTGTGGCGCTGTCTCATCTTCTCGTG 355  
QY 179 GlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198  
Db 356 GGCTCGTGTCTCATGGTTTGGCGTCACTACGCGATGCTCATGGCGGCGGCTCTGTG 415  
QY 199 GluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSer 218  
Db 416 GCGGAGTCTGGTGTGGGCTACGGGGCATGATCGGCCGCTGTACACGGCCGAGATCTCG 475  
QY 219 ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238  
Db 476 CCTGCGCGCTCCCGTGGCTTCTGACCACTTCCCGAGAGGTTCATCAACATCGGCATC 535  
QY 239 PheLeuAlaTyrLeu-----LeuGlyMetPheIleProTrpArg 251  
Db 536 CTGCTTGGCTAGCTGTCCAACTTCGCTTCGCGCGCTCCGCTCCACTCCGCTCGCTGGCGC 595  
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271  
Db 596 GTCATGCTGCCATTTGGCGAGTCTCGTCCGCGCTCTCGCGCTCTCTGTGTCTGCAATG 655  
QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291  
Db 656 CCCGAGTCTGCTCGGTGGTGTGTTGAAGGCGCTCGCGAGCCGCGCCAGCGGCTGTGTA 715  
QY 292 Gln----- 292  
Db 716 GAGAACCTCTGCCACGCCAGAGGCGCGCGAGCGCTGGCGGATCATCAAGCGCGG 775  
QY 293 -----ValLeuArgGlyPheGluThrAspIleThrThr----- 303  
Db 776 GCGGGGATTCGGAAGGCCCTCGACGGGAGCTAGTCACCGTACCGCGCAAGGACGAAGGC 835  
QY 304 -----GluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIle 321  
Db 836 GCGGTGAGTTGCGAGTGTGGAAGAAGCTCATCTGTCCCGACCCCGCTGTC----- 889  
QY 322 SerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeu 341  
Db 890 -----CGACGATCTGCTCTCGGCGGCTGCTC----- 919  
QY 342 LeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePhe 361  
Db 920 CACTTCTTCCAGCAGGCTTCTGCGAGCACTCCGTCAGTACAGCCGCCCTGTTTC 979  
QY 362 LysAlaAlaGlyValThrAsnSerAspLeu-----AlaThrCysSerLeuGlyAla 378  
Db 980 AAGAGCGGGGATCACCGACGACACACAACTCTCTGGGCGTTCACCTGCGCGGCGGCGG 1039  
QY 379 IleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIle 398  
Db 1040 ACCAAGAGTCTTCTCATCTCTGGTGGCCACGTTCTGCTGGACCGCGCGGCGCTCGGCT 1099  
QY 399 LeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhe 418  
Db 1100 CTGCTGTGATCAGACACGGCGGATGATTGTCTCGCTCATCTCGCTCGGCGGCGGCTC 1159  
QY 419 PheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSer 438

```

Db      1160 ACCGTCCGGGGGATCACCCGAGACACCAAGGTGCGTGGGCCGTC-----GCCCTGTGC 1213
Qy      439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTriP 458
Db      1214 ATCGCGTCAACCCCTGCTACATCGCCCTTCTTCATCGCCCTCGGGCCCATCACGGGC 1273
Qy      459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr 478
Db      1274 GTGTACACCTCGGAAATATTCGCGCTCAGGTGCGCGCTGGGCTTCGCGGTGGGTG 1333
Qy      479 LeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr--ThrAsnLeuMetLeuThr 497
Db      1334 GCAGGCAACCGGTGACACGCGGTCTATCCATGACCTTCCTGTCCTCCCTTCCAGGCC 1393
Qy      498 TrpSerValGlyGlyThrPheLeuSerTrpMetValValSerAlaPheThrIleValPhe 517
Db      1394 ATCACCATCGCGGCGAGCTTCTCTCTACTCCGGCATCGCGGTCGCTGGGTTC 1453
Qy      518 ValValLeuTrpValProGluThrLysGly 527
Db      1454 TTCTTACGTGCTCCCGGAGACACGGGC 1483

```

# RESULT 9

```

US-09-894-927B-8
; Sequence 8, Application US/09894927B
; Patent No. 6632924
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; FILE REFERENCE: Targeting of GLUT4
; CURRENT APPLICATION NUMBER: US/09/894,927B
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/591,025
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: US 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: US 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
; NAME/KEY: CDS
; LOCATION: (1)...(2592)
US-09-894-927B-8

Alignment Scores:
Pred. No.:      5,7e-47      Length:      2592
Score:          524.00      Matches:    131
Percent Similarity: 52.33%      Conservative: 82
Best Local Similarity: 32.19%      Mismatches: 166
Query Match:      18.86%      Indels:     28
DB:               4         Gaps:        7

US-10-051-909-36 (1-553) x US-09-894-927B-8 (1-2592)

Qy      133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
Db      562 CTCTCCGTGGCCATCTTTCC-----GTGGCGGCATGATTTCC 600
Qy      153 AlaIleAlaSerGlyGlnMetAlaGluTyriLeGlyArgLysGlySerLeuMetIleAla 172
Db      601 TCCTTCCTCATTTGATATCTCTCAGTGGCTGGAGAGAAAGGGCCATGCTGGTCAAC 660
Qy      173 AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSer----- 190

```

```

Db      661 AATGCTCTGGCGGTGCTGGGGGCGACGCTCATGGGCTGGCCAAACGCTGCTGCTCTCTAT 720
Qy      191 ---PheLeuTyriMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyriThr 209
Db      721 GAAATGCTCATCTTGAGCGATTCCTCATTTGGGCGCTACTACAGGCTGACATCAGGGCTG 780
Qy      210 ValProValTyriLeaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
Db      781 GTGCCCATGTACGTGGGGGAGATTGCTCCACTCACCTGCGGGGCGCCCTTGGGAGCGCTC 840
Qy      230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyri-----Leu 243
Db      841 AACCAACTGGCCATTTTATCGGCATTTCTGATCGCCAGGTCGTGGGCTGGAGTTCCTC 900
Qy      244 LeuGlyMetPheIleProTriPArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db      901 CTGGGCACTGCCAGCGCTGTGGCCACTGCTCTGGGCTCACAGTGTACCTGCGCCCTCTG 960
Qy      264 LeuIleProGlyLeuPhePheIleProGluSerProArgTriPLeu---AlaLysMetAsn 282
Db      961 CAGCTGCTCTGCTGCCCTTCTGTCCGAGAGCGCCGCTACCTTACATCATCCAGAT 1020
Qy      283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAsp 300
Db      1021 CTCGAGGGGCGCTGCCAGAAAGAGTCTGAAGCGCTCGACGGCTGGGCGCCAGTGTTC 1080
Qy      301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThr 320
Db      1081 GTCTGCTGCTGAGTGAAGGATGAGAGCGGAAGCTG-----GAGCGTGAGCGCCACTG 1134
Qy      321 IleSerPheGlnGluLeuAsnGlnLysTyriArgThrProLeuLeuLeuGlyIleGly 340
Db      1135 TCCTGCTCCAGCTCTCTGGCGCGCTACCCACCGCAGCCCTGATCATTTGCGGCTG 1194
Qy      341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyriAlaSerSerIle 360
Db      1195 CTGCACTGAGCCAGCAGCTCTCTGGCATCAATGTGTTTCTATTATTGCGACGATC 1254
Qy      361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380
Db      1255 TTCGACACAGCAGGGGTAGCCAGCTGCTATGCCACCATAGGAGCTGGTGGTCAAC 1314
Qy      381 ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu 400
Db      1315 ACAGTCTTCACCTTGTGCTCGGTGTTGTTGGTGGAGCGGCGGGCGCGGACGCTCCAT 1374
Qy      401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPheLeu 420
Db      1375 CTCTGGGCGCTCGCGGCGATGTGTGGCTGTGCCATCTCATGACTGTGGCTCTGCTCCTG 1434
Qy      421 LysAspAsnIleSerGlnAspSerAsnSerTyriTyriIleLeuThrMetIleSerLeuVal 440
Db      1435 CTGGAGCGAGTTCCA-----GCCATGAGCTACGTTCTCATTTGTG 1473
Qy      441 GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTriPLeuMet 460
Db      1474 GCCATCTTTGGCTTCGTGGCATTTTGTAGATTGGCCCTGGCCCATTCCTTGGTTTCATC 1533
Qy      461 MetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAla 480
Db      1534 GTGGCGAGCTCTTCAGCCAGGAGACCCCGCGGAGCCATGGCTGTGGCTGTGTTCTCC 1593
Qy      481 AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal 500
Db      1594 AACTGACGAGCAACTTCATCATTTGGCATGGGTTCAGTAGTGTTCGGAGGCTATGGGG 1653
Qy      501 GlyGlyThrPheLeuSerTyriMetValValSerAlaPheThrIleValPheValLeu 520
Db      1654 CCCTACGTCTCTCTTCTATTGCGGTCTCTCTGCGGCTCTTCTCATCTTTCACCTCTTA 1713
Qy      521 TrpValProGluThrLysGly 527

```

```

Db      1714 AGAGTACCTGAACATCGAGGC 1734
RESULT 10
US-09-643-597-135
; Sequence 135, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, TongTong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yashir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643.597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-135

Alignment Scores:
Pred. No.: 7,68e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 3 Gaps: 12

US-10-051-909-36 (1-553) x US-09-643-597-135 (1-2856)
QY      27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1 46
Db      14 CCGAGTGAGCAGCCGAGGAGCAGGAGACCAACGAGCGGGGTCGAGTCGAGTCGAG 73
QY      46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
Db      74 TGGAGTCCCGGACCGGAGCAGCAGCCTGAGCGGAGAGCGCGCTCGCAGCGCGCTCG 133
QY      63 gArgProLeuLeuAnThrGlySerTyrTyrArgMetSerSerArgGlnSerSerPheAl 83
Db      134 CCACCGCGTACCCGG-----CGAGCAGAGCCACAGCGCGCGCTGC 178
QY      83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCyth 103
Db      179 CATGGAGCCACGACGACGAGAGCTGACGGGTGCG-----CTCATGCTGGGTGT 226
QY      103 rLeuileValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProth 123
Db      227 GCGAGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 286
QY      123 rGlnAspAlaMetValArgAspLeuAsn----- 132
Db      287 CCAGAAGGTGATCGAGGAGTCTCTACAAACAGACATGGGTCCACCGCTATGGGAGAGCAT 346
QY      133 -----LeuSerIleSerGluPheSerAlaPheG1 142
Db      347 CTGCGCCACCGCTCACCACGCTCTGGTCCCTCTCATGGGCCATCTTCT 398
QY      142 ySerLeuSerAnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
Db      399 -----GTTGGGGGATGATGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 445

```

```

QY      162 rIleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182
Db      446 CTTTGGCGCGGGAATTCATGCTGATGATGAACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 505
QY      182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1 199
Db      506 CATGGGCTTCTCGAAACTGGCAAGTCTTTTGAGATGCTGATCTGGGCGGCTTCATCAT 565
QY      199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluLeuSerPr 219
Db      566 CGGTGTGTACTGCGCGCTGACACACAGGCTTCGTGCCCATGTATGTGGGTGAAGTGTACC 625
QY      219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239
Db      626 CACAGCCTTTCTGTGGGCGCTTGGGCGACCTTGACACAGCTGGGCATCTGCTCGGCATCT 685
QY      239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTyrArgLeuLe 253
Db      686 CATGCCCGAGGTTCGGCTGGGACTCATCATGGGCAACAGGACCTGTGGCCCTGCT 745
QY      253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIleProG1 273
Db      746 GCTGAGCATCATCTTCATCCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
QY      273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291
Db      806 GAGTCCCGCTTCTCTGCTC-----ATCAACCGCAACAGGAGAACCGGCCCAAGAGTGTGCT 862
QY      291 uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311
Db      863 AAAGAAGCTGCGCGG--ACAGCTGACGTGACCATGACCTGCAGGAGATGAAGAAGA 919
QY      311 aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeu--AsnGlnLysLy 330
Db      920 GAGTCGGCAGATGATCGGGAGAGAGAGTCAACCTCTCGAGCTGTTCGCTCCCGCG 979
QY      330 sTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyI1 350
Db      980 CTACCGCAGCCCATCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1039
QY      350 aAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAnSerAs 370
Db      1040 CAACGCTCTCTTATTACTCCAGAGCATCTTCGAGAGCGGGGGTGCAGACGCTGT 1099
QY      370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390
Db      1100 GTATGCCACCATTTGGCTCCGCTATCGTCAACACAGCGCCTTCACCTGCTGCTGCTGTTGT 1159
QY      390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410
Db      1160 GGTGGAGCGAGCAGCGCGGACCTTGACCTCATAGGCTCTGCTGCGCATGGCGGTTG 1219
QY      410 sLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430
Db      1220 TGCCATCTCATGACCATCGCGTAGCACCTGCTGGAGCAGCTACCCCTGG----- 1268
QY      430 rTyrTyrIleLeuMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450
Db      1269 -----ATGCTCTATCTGAGCATCGTGCCATCTTTGGCTTTTGGCTTCITTTGA 1318
QY      450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy 470
Db      1319 AGTGGGTCTCTGGCCCATCCCATGTTTCATCGTGGCTGAACCTCTTCAGCCAGGCTCCACG 1378
QY      470 sLeuLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490
Db      1379 TCCAGCTGCCATTTGCGTTCGAGGCTTCTCCAACTGGAGCTCAAAATTTTCATTTGGGCAT 1438
QY      490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa 510
Db      1439 GTGCTTCCAGTATGGGAGCAACTGTGTGGTCCCTACGTCTTTCATCTTTCATCTGCTGT 1498
QY      510 lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly 527

```

[illegible]

US-09-542-615A-135





```

; APPLICANT: Pan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-135

Alignment Scores:
Pred. No.: 7,68e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-606-421B-135 (1-2856)
QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySerAsnArg1 46
DB 14 CCGAGTGACGACGCCAGGAGGAGGAGAGCCAAACGAGCGGGGTCGGAGTCAGAGTCGAG 73
QY 46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
DB 74 TGGAGTCCCCCGGACCGGAGCGAGCGAGCTGAGCGGGAGAGCGCGCTCGACGCCCGCTCG 133
QY 63 gArgProLeuLeuAanThrGlySerTrpTyArgMetSerSerArgGlnSerSerPheAl 83
DB 134 CCACCGCGCTACCGG-----CGACCGAGAGCCACGAGCGAGCGCTCG 178
QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCystH 103
DB 179 CATGGAGCGCAGCAGCAAGAAGCTGACGGTGC-----CTCATGTGCTGT 226
QY 103 rLeuileValAlaLeuGlyProileGlnPheGlyPheThrSerGlyPheSerSerProth 123
DB 227 GGGAGGAGCAGTGTGCTGGCTCCCTGCGAGTTGGCTACACACATGGAGTCATCAATGCC 286
QY 123 rGlnAspAlaMetValArgAspLeuAan----- 132
DB 287 CCAGAAGTGATCGAGGAGTTCTACACACAGACATGGTCCACCGTATGGGAGAGCAT 346
QY 133 -----LeuSerileSerGluPheSerAlaPheG1 142
DB 347 CTGCGCCACCCAGCTCACCACGCTCGTCCCTCTCAGTGGCCATCTTTCT----- 398
QY 142 ySerLeuSerAanValGlyMetValGlyAlaileAlaSerGlyGlnMetAlaGluTy 162
DB 399 -----GTTGGGGGATGATGCTCTCTCTGTGGGCTTTTCGTTAACCG 445
QY 162 rileGlyArgGlySerLeuMetileAlaAlaileProAanileileGlyTrpLeuAl 182
DB 446 CTTTGGCCCGGGAATTCATGCTGATGATGAACCTGTGGCTTCGTGTCGCGCTGT 505
QY 182 alleSerPheAla-----LysAspAlaSerPheLeuTyMetGlyArgLeuLeuG1 199
DB 506 CATGGGCTTCGMAACTGGGCAAGTCCCTTGAGATGCTGATCCTGGGCGCTTCATCAT 565
QY 199 uGlyPheGlyValGlyileileSerTyThrProValTyThrileAlaGluileSerPr 219

```

```

DB 566 CGGTGTGTACTGGGCGCTGACCACAGGCTTCGTGCCCATGTATGTGGTGAAGTGTCAACC 625
QY 219 oGlnAanMetArgGlyAlaLeuGlySerValAanGlnLeuSerValThrPheGlylePh 239
DB 626 CACAGCCTTTCTGTGGGCGCTGGGACCTCGTGCACAGCTGGGCATCGTCGTCGCGCATCT 685
QY 239 eLeuAlaTyLeuLeuGlyMetPheile-----ProTrpArgileLe 253
DB 686 CATCGCCCGAGGTTCGGCCTGGACTCCATCATATGGGCAACAAGAGACTGTGGCCCTGTCT 745
QY 253 uAlaValileGlyAlaLeuMetLeuileProGlyLeuPhePheileProGl 273
DB 746 GCTGAGCATCATCTTCATCCCGGCTGTGTCGTCATCGTCTGCCCTTCGTCGCCGA 805
QY 273 uSerProArgTrpLeuAlaLeuMetAanLeuThrGluAap-----CyeGluThrSerLe 291
DB 806 GAGTCCCCCTTCCTGCTC---ATCAACCGCAACGAGGAGAACCGGCGCAAGAGTGTGT 862
QY 291 uGlnValleuArgGlyPheGluThrAspPheThrGluValAanAspPheileArgAl 311
DB 863 AAAGAAGCTGCGCGG---ACAGCTGACGTGACCCATGACCTGCAGAGAGATGAAGAAGA 919
QY 311 aValAlaSerSerSerLysArgThrThrileSerPheGlnGluLeu---AanGlnlyLe 330
DB 920 GAGTCGGCAGATGATCGGGAGAGAGAGTCCACCATCTCGAGCTGTTCGCTCCCCCG 979
QY 330 sTyArgThrProLeuLeuGlyileGlyLeuLeuValLeuGlnAanLeuSerGlyil 350
DB 980 CTACCGCCAGCCATCTCTCATCGTGTGTGTCAGCTGTCCAGCAGCTGTCTGCGCAT 1039
QY 350 eAanGlyValLeuPheTyArgSerSerilePheLysAlaAlaGlyValThrAanSerAs 370
DB 1040 CAACGCTGTCTTATTATCTCCAGCAGCATCTTCGAGAGCGGGGTCGACGAGCCTGT 1099
QY 370 pLeuAlaThrCysSerLeuGlyAlaileGlnValLeuAlaThrGlyValThrThrTrle 390
DB 1100 GTATGCCACCATTTGGTCCGGTATCGTCAACAGGCGCTTCACCTGTCTGTGCTGTGTGT 1159
QY 390 uLeuAspArgAlaGlyArgGlyleLeuLeuileileSerThrSerGlyMetThrLeuCy 410
DB 1160 GGTGGAGCGAGCAGCGCGGACCTCGCACCTCATAGGCTCGCTGGCATGGCGGTG 1219
QY 410 sLeuLeuAlaValSerValPhePheLeuLysAspAanileSerGlnAspSerAanSe 430
DB 1220 TGCCATATCATGACCATCGCTAGCACTGTGGAGCAGCTACCTCG----- 1268
QY 430 rTyTrileLeuThrMetileSerLeuValGlyileValSerPheValileThrPheSe 450
DB 1269 -----ATGCTCTATCTGAGCATCGTGCCATCTTTGGCTTTGTGGCCTTCCTTGA 1318
QY 450 rPheGlyMetGlyAlaileProTrpLeuMetMetSerGluileLeuProValSerlely 470
DB 1319 AGTGGGTCTGGCCCATCCCATGGTTTCATGTGGTGAACCTCTTCAGCCAGGCTCCACG 1378
QY 470 sSerLeuGlyGlySerileAlaThrLeuAlaAanTrpLeuThrSerPheAlaileThMc 490
DB 1379 TCCAGCTGCCATTCGGCTTGCAGGCTTCCTCAACTGGACCTCAAAATTCATTTGGG 1438
QY 490 tThrThrAanLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyMetValVa 510
DB 1439 GTGCTTCCAGTATGGAGCAACTGTGTGGTCCCTACGCTTCATCATCTTCACTGTGCT 1498
QY 510 lSerAlaPheThrileValPheValValLeuTrpValProGluThrLyeGly 527
DB 1499 CTGGTTCGTCTTCATCTTCACCTACTTCAAGTTCCTGAGACTAAAGGC 1550

RESULT 14
US-09-221-107-135
; Sequence 135, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER

```

```

FILE REFERENCE: 210121.455C2
CURRENT APPLICATION NUMBER: US/09/221,107
CURRENT FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 161
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 135
LENGTH: 2856
TYPE: DNA
ORGANISM: Human
US-09-221-107-135

Alignment Scores:
Pred No.: 7,680-47 Length: 2856
Score: 52350 Matches: 151
Percent Similarity: 55.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-221-107-135 (1-2856)
QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySerAenArgG1 46
DB 14 CCAGGTGACGACCGAGGAGCAGGACCAACACCGCGGGTCCGAGTCAGAGTCGAC 73
QY 46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
DB 74 TGGAGTCCCGGACCGGAGCAGGAGCCTGAGCGGAGAGCGCGCTCGCACGCCCGCTG 133
QY 63 gArgProLeuLeuAsnThrGlySerTyrTyrArgMetSerSerArgGlnSerSerPheAl 83
DB 134 CMCCCGCTACCGCG-----CGAGCCAGGAGCACCAGCGCGAGCGGTGC 178
QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCySth 103
DB 179 CATGGAGCCACGACAGACGCTACCGGTGCG-----CTCATGCTGGCTGT 226
QY 103 rLeuLeuValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProTh 123
DB 227 GCGAGGACGAGTGTGTGCTCCTCGAGTTTGCTACACACTGGAGTCATCAATGCCCC 286
QY 123 rGlnAspAlaMetValArgAspLeuAsn----- 132
DB 287 CCAGAAGGTGATCGAGGAGTCTTACACACAGATGGGTCCACCGCTATGGGAGAGCAT 346
QY 133 -----LeuSerIleSerGluPheSerAlaPheG1 142
DB 347 CTGCGCCACCGCTACACGCTCTGTGCTCCTCTCAGTGGCCATCTTTCT----- 398
QY 142 ySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
DB 399 -----GTTGGGGCATGATGGTCTCTCTCTGTGGGCTTTTTCGTTAAACCG 445
QY 162 rIleGlyArgIysGlySerLeuMetIleAlaIleProAsnIleIleGlyTyrLeuAl 182
DB 446 CTTTGGCGCGGGAATTCAATGTGATGATGAACCTGCTGCGCTTCGTGTCGCCGTGT 505
QY 182 alleSerPheAla-----IysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1 199
DB 506 CATGGGCTTCTCGAAACGCGCAAGTCTTGTAGATGCTGATCTCGGCGCTTTCATCAT 565
QY 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219
DB 566 CGGTGTGTACTGCGGCTTGACACAGCGCTTGTGCCCATGATGTGGGTGAAGTGTACCC 625
QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239
DB 626 CACAGCCTTTCGTGGGCGCTTGGGACACCTGACACCGAGCTGGCATCGTGGCATCTC 685
QY 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTyrArgLeuLeu 253
DB 686 CATGCGCCGAGTGTGCGCTGGAGTCCATCATGATGGGCAACAGGACCTGTGCGCCCTGCT 745

```

---

```

253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273
746 GCTGAGCATATCTTCATCCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
273 uSerProArgThrProLeuAlaIysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291
806 GAGTCCCGCTTCTCTGCTC---ATCAACCGCAACAGAGAGAACCGGCGCAAGAGTGTCT 862
291 uGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleIysArgAl 311
863 AAAGAAGCTGCGCGG---ACAGCTGAGCTGACCATGACCTGCAGGAGATGAAGAAGA 919
311 aValAlaSerSerSerIysArgThrThrIleSerPheGlnGluLeu---AsnGlnIysLy 330
920 GAGTCGGCAGATGATGCGGAGAGAAAGGTCAACATCTCTGGAGCTGTTCCGCTCCCGCC 979
330 sTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIyl 350
980 CTACCGCCAGCCCATCTCATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039
350 eAsnGlyValLeuPheTyrAlaSerSerIlePheIysAlaAlaGlyValThrAsnSerAs 370
1040 CAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAAGCGGCGGTGCAGCAGCTGT 1099
370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTriple 390
1100 GTATGCCACCATGGCTCCGCTATCGTCAACACGCGCTTTCACCTGCTGCTGCTGCTGT 1159
390 uLeuAspArgAlaGlyArgGlyLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410
1160 GGTGGAGCAGACGAGCGCGGACCTGACCTCATAGGCTGCTGCTGCTGCTGCTGCTGCTG 1219
410 sLeuLeuAlaValSerValPhePheLeuIysAspAsnIleSerGlnAspSerAsnSe 430
1220 TGCCATCTACTCATGACCATCGCTAGCTGCTGAGGAGCAGTACCTCG----- 1268
430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450
1269 -----ATGCTCTATCTGAGCATCGTGCCATCTTTGCTGCTGCTGCTGCTGCTGCT 1318
450 rPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleLy 470
1319 AGTGGGCTGCGGCGCCATCCATGCTTCATCGTGGCTGAACCTTTCAGCCAGGCTCCACG 1378
470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThrMe 490
1379 TCCAGCTGCCATTTGCCGTTGCGAGGCTTCTCCAACTGGACCTCAAATTTTCATTGGCG 1438
490 tThrThrAsnLeuMetLeuThrTyrSerValGlyGlyThrPheLeuSerTyrMetValVa 510
1439 GTGCTTCCAGTATGAGGACACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1498
510 lSerAlaPheThrIleValPheValValLeuTyrValProGluThrLysGly 527
1499 CTGGTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1550

RESULT 15
US-09-466-396A-135
; Sequence 135, Application US/09466396A
; Patent No. 6696247
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien

```

US-09-466-396A-135

Alignment Scores:

Pred. No.: 7.68e-47 Length: 2856  
 Score: 523.50 Matches: 151  
 Percent Similarity: 45.91% Conservative: 96  
 Best Local Similarity: 28.07% Mismatches: 229  
 Query Match: 18.84% Indels: 62  
 DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-466-396A-135 (1-2856)

QY	27	ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1	46
DB	14	CCGAGTGAGCACCCAGGGAGGAGGACCAACAGCGGGGTCGGAGTCAGATCGCAG	73
QY	46	YGIYAla-----GlyValaGlyGluGluSerGlySerAspHisAspGlyValLeuAr	63
DB	74	TGGAGTCCCGGACCGGACGAGCGCTGAGCGGGAGAGCGCGCTCGCACCGCGTGC	133
QY	63	gArgProLeuLeuAsnThrGlySerTrpTyArgMetSerSerArgGlnSerPheAl	83
DB	134	CCACCCGCGTACC CGG-----CGCAGCCAGAGCCACCCAGCGCGCTGC	178
QY	83	aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCyth	103
DB	179	CATGGAGCCCGCAGCAGCAAGATGAGCGGTGCG-----CTCATGTGCTGT	226
QY	103	rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProTh	123
DB	227	GGAGAGAGCGTCTTGCTGCCCTGCGATTTGGCTACACACTGGAGTCATCATGCCCC	286
QY	123	rGlnAspAlaMetValArgAspLeuAsn-----	132
DB	287	CCAGAGGTGATCGAGGAGTTCTACACACAGACATGGGTCCACCGCTATGGGAGAGCAT	346
QY	133	-----LeuSerIleSerGluPheSerAlaPheG1	142
DB	347	CCTGCCACCCACCGCTCACCAGCTCTGGTCCCTCTCAGTGGCCATCTTTCT	398
QY	142	ySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy	162
DB	399	-----GTTGGGGCGATGATGTGCTCTCTCTGTGGGCTTTTCGTTAACCG	445
QY	162	rIleGlyArgLyGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAl	182
DB	446	CTTTGGCCGCGGAATTCATGCTGATGATGAACCTGCTGGCCTCGTGTCCCGCTGT	505
QY	182	alleSerPheAla-----LysAspAlaSerPheLeuTyMetGlyArgLeuG1	199
DB	506	CATGGGCTTCTCGAAACTGGGCAAGTCCCTTTGAGATGCTGATCCTGGGCGCTTCATCAT	565
QY	199	uGlyPheGlyValGlyIleIleSerTyThrValProValTyIleAlaGluIleSerPr	219
DB	566	CGGTGTGTACTGGGCGCTGACCACAGCTTCGGTGGCCCATGATGTGGGTGAAGTGCACC	625
QY	219	oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh	239
DB	626	CACAGCTTTCTGTGGGGCCCTGGGACCCCTGCACAGCTGGGCATCGTCTCGGCATCCT	685
QY	239	eLeuAlaTyLeuLeuGlyMetPheIle-----ProTrpArgLeuLe	253
DB	686	CATCGCCAGGTGTTGGGCTGGAGCTCCATCATGGGCAACAGGACCTGTGGCCCTGTCT	745
QY	253	uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1	273
DB	746	GCTGAGCATCATCTTCATCCCGCCCTGCTGCGAGTGCATCGTCTGCCCTTCTGCCCCGA	805
QY	273	uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe	291
DB	806	GAGTCCCGCTTCTGCTC-----ATCAACCGCAACGAGGAGACCGGGCCAGAGTGTGCT	862
QY	291	uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl	311

DB	863	AAAGAAGCTGCGCGGG-----ACAGCTCAGCTGACCCATGACCTCGAGAGATGAAGAAGA	919
QY	311	aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeu----AsnGlnLysLy	330
DB	920	GAGTCCGCGAGATGATCGGGAGAGAAGGTCCACCATCTGGAGCTGTTCGCTCCCCCGC	979
QY	330	sTyArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyI1	350
DB	980	CTACCCGCGAGCCCATCTCATCTGTGTGTGTGTCAGCTGTCCAGCAGCTGTCTGGCAT	1039
QY	350	eAsnGlyValLeuPheTyAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs	370
DB	1040	CAACGCTGCTTCTATTACTCCACGAGCATCTTCGAGAAGCGGGGTGCGAGCGCTGT	1099
QY	370	pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe	390
DB	1100	GTATGCCACCATTTGGTCCGGTATCTCAACACGGCTTCACCTGTGTGTGTGTGT	1159
QY	390	uLeuAspArgAlaGlyArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy	410
DB	1160	GGTGGAGCGAGCAGCGCGGCGGACCTCATAGGCTCGCTGGCATGGCGGTTG	1219
QY	410	sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe	430
DB	1220	TGCCATACTCATGACCATCGCTAGCATCTGTGGAGCAGCTACCTCG-----	1268
QY	430	rTyTyIleLeuLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe	450
DB	1269	-----ATGCTCTATCTGAGCATCTGTGGCCATCTTTGGCTTTTGTGGCCTTCTTGA	1318
QY	450	rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlyIleLeuProValSerIleLy	470
DB	1319	AGTGGGTCTGGCCCATCCCATGCTTCATCGTGGTGAACCTCTTCAGCCAGGTCACG	1378
QY	470	sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe	490
DB	1379	TCAGCTGCCATTTGGCGTTGCGAGGCTTCCCACTGGACCTCAAAATTTTATTGTGGCAT	1438
QY	490	tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyMetValVa	510
DB	1439	GTGCTTCCAGTATGTGGACCAACTGTGTGGTCCCTACGCTTCATCATCTTCACTGTGCT	1498
QY	510	lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly	527
DB	1499	CCTGGTCTGTCTTCTCATCTTCACCTACTTCAAAAGTTCCTGAGACTAAAGGC	1550

Search completed: April 13, 2005, 22:00:59  
 Job time : 264.162 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: April 13, 2005, 18:06:15 ; Search time 702.592 Seconds  
(without alignments)  
4775.020 Million cell updates/sec

Title: US-10-051-909-36  
Perfect score: 2779  
Sequence: 1 PSSSSFRPAGKKKKKNOG.....TIFVLSIORQLWLPECLS 553

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10051909/runat\_13042005\_074039\_14141/app.query.fasta\_1.1678  
-DB=PublishedApplications\_NA\_QFMT=fastap -SUFFIX=rnnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10051909\_@CGN 1 1 953 @runat\_13042005\_074039\_14141  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOC=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/prodata/2/pubpna/US10D\_PUBCOMB.seq\*
- 17: /cgn2\_6/prodata/2/pubpna/US10E\_PUBCOMB.seq\*
- 18: /cgn2\_6/prodata/2/pubpna/US10F\_PUBCOMB.seq\*
- 19: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq\*
- 20: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq\*
- 21: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq\*
- 22: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2777	99.9	1953	13	US-10-051-909-35 Sequence 35, Appl
2	2133	76.8	2003	17	US-10-437-963-61061 Sequence 61061, A
3	2129.5	76.6	1797	17	US-10-425-114-26596 Sequence 26596, A
4	2104	75.7	1875	17	US-10-425-114-26550 Sequence 26550, A
5	1806.5	65.0	2049	17	US-10-425-114-24968 Sequence 24968, A
6	1806.5	65.0	2244	18	US-10-425-115-138759 Sequence 138759, A
7	1797.5	64.7	1459	18	US-10-425-115-35618 Sequence 35618, A
8	1718	61.8	2202	17	US-10-424-599-132422 Sequence 132422, A
9	1374	49.4	1304	18	US-10-437-963-63465 Sequence 63465, A
10	1356	48.8	1167	18	US-10-425-115-47714 Sequence 47714, A
11	1312.5	47.2	1105	18	US-10-425-115-47714 Sequence 47714, A
12	1221.5	44.0	957	17	US-10-425-114-28778 Sequence 28778, A
13	1185	42.6	1283	17	US-10-424-599-131066 Sequence 131066, A
14	1182.5	42.6	1077	18	US-10-767-795-565 Sequence 565, App
15	1139.5	41.0	2105	17	US-10-424-599-132426 Sequence 132426, A
16	1139	41.0	1027	17	US-10-425-114-5566 Sequence 5566, App
17	1081	38.9	1498	18	US-10-425-115-175050 Sequence 175050, A
18	946.5	34.1	990	18	US-10-767-795-566 Sequence 566, App
19	944.5	34.0	1828	18	US-10-739-930-1250 Sequence 1250, App
20	941.5	33.9	785	18	US-10-425-115-140613 Sequence 140613, A
21	920.5	33.1	1473	9	US-09-938-842A-2331 Sequence 2331, App
22	920.5	33.1	1473	11	US-09-938-842A-2331 Sequence 2331, App
23	909.5	32.7	1785	18	US-10-437-963-64291 Sequence 64291, A
24	899	32.3	1359	9	US-09-938-842A-1627 Sequence 1627, App
25	899	32.3	1359	11	US-09-938-842A-1627 Sequence 1627, App
26	893	32.1	1768	17	US-10-310-154-357 Sequence 357, App
27	888.5	32.0	1963	18	US-10-739-930-1789 Sequence 1789, App
28	877.5	31.6	1407	9	US-09-938-842A-1291 Sequence 1291, App
29	877.5	31.6	1407	11	US-09-938-842A-1291 Sequence 1291, App
30	777	28.0	1374	9	US-09-938-842A-1020 Sequence 1020, App
31	777	28.0	1374	11	US-09-938-842A-1020 Sequence 1020, App
32	759	27.3	1635	18	US-10-437-963-4124 Sequence 4124, App
33	733	26.4	1231	17	US-10-424-599-132423 Sequence 132423, A
34	694	25.0	1319	18	US-10-425-115-39803 Sequence 39803, A
35	693.5	25.0	1779	17	US-10-424-599-104365 Sequence 104365, A
36	679.5	24.5	778	17	US-10-425-114-22347 Sequence 22347, A
37	669.5	24.1	2080	16	US-10-168-651-28 Sequence 28, Appl
38	666	24.0	1258	17	US-10-424-599-132425 Sequence 132425, A
39	663	23.9	1473	9	US-09-938-842A-1669 Sequence 1669, App
40	663	23.9	1473	11	US-09-938-842A-1669 Sequence 1669, App
41	658.5	23.7	1862	17	US-10-115-831-137 Sequence 137, App
42	657.5	23.7	600	18	US-10-021-323-14004 Sequence 14004, A
43	635	22.8	1242	18	US-10-425-115-177046 Sequence 177046, A
44	626	22.5	668	18	US-10-425-115-137175 Sequence 137175, A
45	625.5	22.5	890	17	US-10-425-114-127 Sequence 127, App

# ALIGNMENTS

RESULT 1  
US-10-051-909-35  
; Sequence 35, Application US/10051909  
; Publication No. US20020199217A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Helentjaris, Tim  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: B01163 US CIP  
; CURRENT APPLICATION NUMBER: US/10/051,909  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 60/083,044  
; PRIOR FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 35  
; LENGTH: 1953



Qy	254	AlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGlu	273
Db	808	GCTGTAAATAGGAATCTGCCTTGACATGTTGTATACCTGGCCTATTCTTCATTCACGAA	867
Qy	274	SerProArgTrpLeuAlaLysMetAsnLeuThrGluAaspCysGluThrSerLeuGlnVal	293
Db	868	TCCCCAAGATGGTTGGCAAGATGAACATGATCGATGATTTTGGAGACTTCTTTTACAAGTT	927
Qy	294	LeuArgGlyPheGluThrAspIleThrGluValIasnAspIleLysArgAlaValAla	313
Db	928	CTGAGGGGATTTTGAGACTGACATCAGCGCGGAAGTGAATGATATAAAGAGAGCAGTAGCG	987
Qy	314	SerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysLysValArgThr	333
Db	988	TCAGCAACAAAAGACACAGTCCGTTTTCAGAAATTAACCCNAGAAANAATACCGCAC	1047
Qy	334	ProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyVal	353
Db	1048	CCCTTAATCTAGGAATTGGCCTACTTGTACTGCAACAGCTAAGTGGAAATCAATGGAA	1107
Qy	354	LeuPheTyrrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThr	373
Db	1108	TTGTTTATACGAGGTAGCATCTTCAAAGCAGCAGGTCTCAAAACAGTGACTTGGCTACA	1167
Qy	374	CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArg	393
Db	1168	TGTGCACCTGGTCTATCCAGGTTCTTGCTACAGGAGTTACAACTGGTTATTAGACAGA	1227
Qy	394	AlaGlyArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAla	413
Db	1228	GCTGGCGCAGGATCCTCCTTATCATCTCTCTGCTGGGATGACTTAAGCCTCCTTGCA	1287
Qy	414	ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrrTyrrIle	433
Db	1288	GTTGCTGTGTATTTTTTCCCAAGGATAGCATTTCAACAAGATTTCTACATGTACTACACC	1347
Qy	434	LeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet	453
Db	1348	TTAAGTATGATCTCCTTGGTTGCTCTGTGGCTTTTGTATCGCCTTCTCCTTCGGTATG	1407
Qy	454	GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly	473
Db	1408	GGTGCCATTCATGGATCATAAATGTCAGAGATCTCCCGGTAGTATCAAGAGTCTCGCA	1467
Qy	474	GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrAsn	493
Db	1468	GSAAGCTTTCGACGCTCGCCAACTGGCTTACATCCTTTGGAATAACAATGACAGCAAC	1527
Qy	494	LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrrMetValValSerAlaPhe	513
Db	1528	TTGATGCTTAGCTGGAGTGCTGGAGGACCTTGTGTCTTACATGGTCGTGAGTGCCTTC	1587
Qy	514	ThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp	533
Db	1588	ACCCTCGTGTGCTCATCCTTTGGTGCCAGAGACAAAGGAAG-AACTCTCGAAGAGAT	1646
Qy	534	ThrIlePheValSerLeuSerIle	541
Db	1647	ACAATGGTCTTCGCTGAGCCCT	1670

RESULT 3

US-10-425-114-26596

; Sequence 26596, Application US/10425114

; Publication No. US2004003488A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 26596  
LENGTH: 1797  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURES:  
OTHER INFORMATION: Clone ID: LIB4371-017-F11\_FLI  
US-10-425-114-26596

Alignment Scores:  
Pred. No.: 1,026-202 Length: 1797  
Score: 2129.50 Matches: 424  
Percent Similarity: 91.91% Conservative: 42  
Best Local Similarity: 83.63% Mismatches: 36  
Query Match: 76.63% Indels: 6  
DB: 17 Gaps: 4

US-10-051-909-36 (1-553) x US-10-425-114-26596 (1-1797)

QY 40 GlyGlyGlySerAsnArg---GlyGlyAlaGlyAlaGlyGluGlySerGlySerAspHis 58  
DB 89 GCGCGCGCCATCAACAGCGCGCGCGCTCGCGCGATGAGCGCGAAGCGACTAC 148  
QY 59 Asp-----GlyValLeuArgProLeuLeu---AsnThrGlySerTyrArgMet 75  
DB 149 GAGCGCGCGCGCGATCGCGAAGCGCTGCTGATGCACACGGGAGCTGGTACAGGATG 208  
QY 76 SerSerArgGlnSerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArgGlu 94  
DB 209 GGGTCGCGCGCGGAGCGCTCACCGCGCGGAGACTCGTCCATGGCGATCCCTACCGCAG 268  
QY 95 SerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProIleGlnPheGly 114  
DB 269 TCCACAGCTCTCGCGCTTCTCTGCAAGCTCATCGFCGCGCTCGCGCCCATCCAGTTCGGA 328  
QY 115 PheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSer 134  
DB 329 TTCACGGGGGGCTTCTCTCCCGCAGCAGGACGCATCATCCGAGACCTCGACCTCACC 388  
QY 135 IleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIle 154  
DB 389 CTCCTCGAGTCTCGGTGTTCCGATCGTGTCCAAAGCTCGCGCCATGTTGGGGCGATT 448  
QY 155 AlaSerGlyClnMetAlaGluTyrIleGlyArgGlySerLeuMetIleAlaAlaIle 174  
DB 449 GCAGTGGTGCAGATGGCGGAGTACATGGCGCAAGGCTCATTCGATGATTCGTCGATTT 508  
QY 175 ProAsnIleGlyTyrIleAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMet 194  
DB 509 CCAACATCATTTGGTGGCTTCCCATCTCTTTGCAAAAGACTCATCGTTTCITTTATATG 568  
QY 195 GlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIle 214  
DB 569 GCACGATTCTCGAGGGGTTGGTGTGGTGTGTCATCTTATACGGTGCACAGTTTACATA 628  
QY 215 AlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerVal 234  
DB 629 GCAGAAATATACCTCAAAACATGAGAGTGCTCTTGGCTCAGTCAATCAGTTATCTGTA 688  
QY 235 ThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAla 254  
DB 689 ACCGTGGTATATTTGGCATATTTGCTCGCATGTTTGTTCCTTGGAGGCTTCTTGTCT 748  
QY 255 ValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGlySer 274  
DB 749 GTAAATAGGAATCTTGCCTTGCATGTGTGATGTTGACCTGGCCCTATCTTCTTCAGATCC 808  
QY 275 ProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeu 294  
DB 809 CCAAGATGGTGGCAAAAGATGAACATGATGGATGATTTTGAGACTTCTTTTCAAGATTCTG 869

QY 295 ArgGlyPheGluThrAspIleThrThrGluValAsnAspIleIysArgAlaValAlaSer 314  
DB 869 AGGGATTTTGAGACTGACATCAGCGCGAAGTGAATATATAAGAGAGCGATGGCTCA 928  
QY 315 SerSerIysArgThrThrIleSerPheGlnGluLeuAsnGlnIlyIysTyrArgThrPro 334  
DB 929 GCAAAACAAAGGACACAGATCCCGTTTCAAGAATTAAACCCAGAGAAATACCCGACACC 988  
QY 335 LeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeu 354  
DB 989 CTAATACCTAGGAATTGGCTACTTGTACTGCAACAGCTAAGTGAATCAATGAATATTG 1048  
QY 355 PheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCys 374  
DB 1049 TTTTATGAGGTAGATCACTTCAAGAGCAGAGGCTCTCAAAACAGTGACTTGGCTACATGT 1108  
QY 375 SerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTyrLeuLeuAspArgAla 394  
DB 1109 GCACCTGGTGCTATCCAGGTTCTTGTACAGAGGTTACAACCTGGTTATTAGACAGCT 1168  
QY 395 GlyArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaVal 414  
DB 1169 GCGCGACGAGTCTCTTATCATCTCTGCTGGGATGACTCTAAGCCCTCTTGAGTT 1228  
QY 415 SerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu 434  
DB 1229 GCTGTGTGTTATTTTCTCAAGGATAGCATTTCACAAGATTCTCACATGTACTACACTTA 1288  
QY 435 ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGly 454  
DB 1289 AGTATGATCTCTTGGTGTCTTGTGCTTTTGTAACTCGCTTCTCTCTCGGTATGGGT 1348  
QY 455 AlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleIysSerLeuGlyGly 474  
DB 1349 GCCATTCATGATGATCAATGTCAGAGATCTCTCCGCTAGTATCAAGAGTCTCGCAGGA 1408  
QY 475 SerIleAlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThrMetThrThrAsnLeu 494  
DB 1409 AGCTTTGGAGCGCTCGCACTGCTTACATCTTTCGATATACATGACAGCAAACTTG 1468  
QY 495 MetLeuThrTyrSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThr 514  
DB 1469 ATGCTTAGCTGGAGTGTGGAGGAGACCTTTGTGCTTACATGGTGGTGGTCTTCCACC 1528  
QY 515 IleValPheValValLeuTyrValProGluThrIysGly\*\*\*AsnSerArgGlyAspThr 534  
DB 1529 CTCGTGTTTCGTCACTCTTTGGTGTCCAGAGACAAAGGAAG-AACTCTCGAAGAGATACA 1587  
QY 535 IlePheValSerLeuSerIle 541  
DB 1588 ATGGTCTCTCCGCTGAGGCTT 1608

RESULT 4  
US-10-425-114-26550  
Sequence 26550, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Kovalic, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 26550  
LENGTH: 1875  
TYPE: DNA



[illegible]

;  
; APPLICANT: Cao, Yongwei  
;  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
;  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement



## Alignment Scores:

Pred. No.: 3,15e-170 Length: 2244  
 Score: 1806.50 Matches: 366  
 Percent Similarity: 81.04% Conservations: 70  
 Best Local Similarity: 68.03% Mismatches: 79  
 Query Match: 65.01% Indels: 26  
 DB: 18 Gaps: 4

US-10-051-909-36 (1-553) x US-10-425-115-138759 (1-2244)

```

Qy 34 LeuArgThrArgValMetGlyGlyGlySerAsnArgGlyAlaGlyAlaGlyGluGlu 53
    |||||
Db 359 CTAGGATGAGTCTCGGACACAGAGAGTGGCGGGAGATG-GGGGAGAGC-TC 416
    |||||

Qy 54 SerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTrpTyr 73
    |||||
Db 417 TCGGCTCCGAC-----CTGCGGAAGCGGCTCATCAACCGGAGCTGGTAC 464
    |||||

Qy 74 Arg-----MetSerSerArgGlnSerSerPheAlaProGlyThr 86
    |||||
Db 465 CGATGCGCGCGCGGTGGCTGATGGCTCGCGGCAATCTAGCCCTCATGGAGCGATTG 524
    |||||

Qy 87 SerSerMetAlaVal---LeuArgGluSerHisValSerAlaPheLeuCysThrLeuLe 105
    |||||
Db 525 GGCCTCTCTCGGTCTCTCTCGGAGCTGCTATCTCGGAGAGCTCTCGACGCTCAT 584
    |||||

Qy 106 ValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProThrGlnAsp 125
    |||||
Db 585 GTCGCGTAGTCCCATCCAGTTTCGGTTTCACATGCGGCTACTCTCGCCACCGAGGAC 644
    |||||

Qy 126 AlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSer 145
    |||||
Db 645 GCCATCATGCTGATCTCGGCCCTCTCCCTCTGAGTTCTCCCTTCGCTTCATTATCT 704
    |||||

Qy 146 AsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArg 165
    |||||
Db 705 AATGTAGGGCGAGTGTAGGCGGCATCTCCAGTGGGCACTTCGAGAGTATATCGGCCGC 764
    |||||

Qy 166 LysGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPhe 185
    |||||
Db 765 AAGGGGTCTCTCATGCTCGCGATTCCAAACATAAATTTGGTGGCTCGCGATATCATTC 824
    |||||

Qy 186 AlaIysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyLe 205
    |||||
Db 825 GCAAAAGATTCTCTTTCTTTTATGGTTCGGCTGCTAGAGGATTTGGAGTCGGTGA 884
    |||||

Qy 206 IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
    |||||
Db 885 ATATCGTATACGATACCGGTTTATATTCAGAAATCGCTCTCAGGATCAGAGGGAGCT 944
    |||||

Qy 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly 245
    |||||
Db 945 CTGGTCTCTCAATCAGCTCTCGTCACGATTGGTATATGCTTGCTTACTTGTGTGGC 1004
    |||||

Qy 246 MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuLe 265
    |||||
Db 1005 ATGTTTGTTCCTCGGAAATCTTGTGTCTAGGCATTTTACCTTGTTCAATCTCGATT 1064
    |||||

Qy 266 ProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaIysMetAsnLeuThrGlu 285
    |||||
Db 1065 CTTGGACTGTCTTGTGGCTGAATCCCCAAGTGGTGGCAAAATGGGGAAGATGGAG 1124
    |||||

Qy 286 AspCysGlyThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluVal 305
    |||||
Db 1125 GATTTTGAATATTCAATGTCAGTTCTGAGAGGATTTTCAGACATATCAGACGAGATGA 1184
    |||||

Qy 306 AsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGlu 325
    |||||
Db 1185 AATGAAATAAGAGATCATTAGCATCATCGAGGAGGAGGACCAACCATTAAGTTTCGTGAT 1244
    |||||

Qy 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGln 345
    |||||
  
```

```

Db 1245 ATCAACAGAGAGATACAGTGTCCCTTGTGTAGTAATCGGTCTCTCTTGTCTGCAG 1304
Qy 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
    |||||
Db 1305 CAGCTAAGTGGTGTCAATGCGATCTATTTTATGCTGCGAGCATCTTCAAAGCTGTGCT 1364
    |||||

Qy 366 ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly 385
    |||||
Db 1365 ATTACAAACAGTAATCTAGCAACATTTGGTTTAGGGGCTGTTCAGGTGATTGTCTACTGGA 1424
    |||||

Qy 386 ValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuLeuIleIleSerThrSer 405
    |||||
Db 1425 GTGACAACCTGGTTGACTGACAAAGCTGGTTCGAAGGCTTCTTCTCATTTATTCACCACA 1484
    |||||

Qy 406 GlyMetThrLeuLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSer 425
    |||||
Db 1485 GGAATGGTCAATCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1544
    |||||

Qy 426 GlnAspSerAsnSerTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe 445
    |||||
Db 1545 GCTGGTTCGCACCTTACTCTGTAATGAGTATGCTTTCACCTGGCTGGACTTGTGGCATTT 1604
    |||||

Qy 446 ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeu 465
    |||||
Db 1605 GTGATTGCAATTTCTCTGGCTGGGAGCGATTCCGTGGATCATTTATGCTGAGATCCTT 1664
    |||||

Qy 466 ProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer 485
    |||||
Db 1665 CTTGTTAAACATCAAGAGCTTGTCTGGAAGTGTTCGCAACCTGCGCAACCTGCTGACAGCA 1724
    |||||

Qy 486 PheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeu 505
    |||||
Db 1725 TGGGCCATACATAATGACGCAAGCCGTGTTGAACCTGGAGCAGTGGAGGAACCTTTGCT 1784
    |||||

Qy 506 SerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThr 525
    |||||
Db 1785 ATCTAGCCGCTGCTGTCTACCATGGCCCTCATTTTCTGTGCTTGTGGTGGCTGAGACC 1844
    |||||

Qy 526 LysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeu 545
    |||||
Db 1845 AAGGGAAG-AACGCTAGAGAAATCGCTTCTCATTCGCTGACACGTCGTCTATGATCTA 1903
    |||||

Qy 546 -GlnTrp-----LeuProGluCysLeu 552
    |||||
Db 1904 GGTATGGGAAGCCACACACCATGTCATGTCATGTCGCGCGGTGGCTT 1955
    |||||
  
```

## RESULT 7

US-10-425-115-35618  
 ; Sequence 35618, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 35618  
 ; LENGTH: 1459  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRF4577\_132481C.1  
 ; US-10-425-115-35618

Alignment Scores:  
 Pred. No.: 1,32e-169 Length: 1459  
 Score: 1797.50 Matches: 372  
 Percent Similarity: 98.15% Conservative: 0

Best Local Similarity: 98.15% Mismatches: 7  
 Query Match: 64.68% Indels: 3  
 DB: 18 Gaps: 0

US-10-051-909-36 (1-553) x US-10-425-115-35618 (1-1459)

Qy	175	ProAenileleGlyTripleuAlaSerPheLeuTyMet	194
Db	3	CCCAATTCATCGGTGGCTTGGATCCCTTGGCAAGA-GCCTCATTTATATATG	61
Qy	195	GlyArgLeuLeuGluGlyPheGlyValGlyLeuSerTyThrValProValTyLe	214
Db	62	GGACGATTCCTTGAAGGGTTGGTTCGGATCATATCTACACGGTACCGGTATACATA	121
Qy	215	AlaGluLeuSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerVal	234
Db	122	GCAGATATCTCTCGAACAATGAGGGAGCTTGGTCTGTGAACAGTCTCTGTG	181
Qy	235	ThrPheGlyLeuPheLeuAlaTyThrLeuGlyMetPheLeuProTyArgLeuAla	254
Db	182	ACCTTTGGCATATCTTGGCCATTTTCTCGCATGTTTATCTCTGGAGACTTCTTGT	241
Qy	255	ValileGlyAlaLeuProCysTyThrMetLeuLeuProGlyLeuPheLeuProGluSer	274
Db	242	GTGTTGGAGCCCTTGGCTGCAATGTTGATTCCTGGGCTATCTTCATTTCCAGAACT	301
Qy	275	ProArgTripleuAlaLeuMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeu	294
Db	302	CCAGATGGCTTGCACAAAGATGATTTGATGGAGATTTGAGACGCTCCCTACAAAGTGTG	361
Qy	295	ArgGlyPheGluThrAspLeuThrThrGluValAsnAspLeuLeuValAlaSer	314
Db	362	AGGGGTTTGGGACTGCATCAACACAGAGTGAATGATATATAAAGAGGCGAGTGCATCA	421
Qy	315	SerSerArgThrThrLeuSerPheGlnGluAsnGlnLeuTyThrArgThrPro	334
Db	422	TCAAGTA--AGACCAACATCAGTTTCAAGAAATTAACCAAAAGAAATACCGACGCCA	479
Qy	335	LeuLeuGlyLeuGlyLeuValLeuGlnAsnLeuSerGlyLeuAsnGlyValLeu	354
Db	480	CTACTTCTAGGATTTGGCTACTTGTACTGCAAAATCTTAGTGAATCAACGGTGTACTG	539
Qy	355	PheTyAlaSerSerilePheTyAlaGlyValThrAsnSerAspLeuAlaThrCys	374
Db	540	TTTTATCAAGTAGCATCTCAAGCTGCAGGGGTTTACAAACACAGCACTTGGCCACCTGT	599
Qy	375	SerLeuGlyAlaLeuGlnValLeuAlaThrGlyValThrThrTripleuLeuAspArgAla	394
Db	600	TCACTTGGTCTATTCAAGGTTCTTGGTCTTACTGAGTTTACAACTGGCTGTAGACGAGCT	659
Qy	395	GlyArgArgileLeuLeuLeuLeuLeuSerThrSerGlyMetThrLeuCysLeuLeuAlaVal	414
Db	660	GGACGACGCATCTTCTCATTTTCTACCTCTGGCATGACTTATGCTTCTTGGCCGT	719
Qy	415	SerValValPhePheLeuTyAspAsnLeuSerGlnAspSerAsnSerTyThrileLeu	434
Db	720	TCGTGTGATTTTTCTCAAGGATAACATTTACAGAGTTCTAACTCATCTACTATCTTA	779
Qy	435	ThrMetileSerLeuValGlyLeuValSerPheValileThrPheSerPheGlyMetGly	454
Db	780	ACATGATCTCCCTTGTGGTATTTGTCTTTTGTCTTATCATCTTCTCGTTTGGTATGGT	839
Qy	455	AlaileProTripleuMetSerGluileLeuProValSerileysSerLeuGlyGly	474
Db	840	GCCATTCATGGCTCATGATGTCGTGAGATCCTCCCGGTTAGCATCAAGAGCCTTGGCCGA	899
Qy	475	SerileAlaThrLeuAlaAsnTripleuThrSerPheAlaileThrMetThrThrAsnLeu	494
Db	900	AGCATCGCAACACTGGCCAACTGGCTGCATCTCTCGCCATAACAATGACGACGAACTTG	959
Qy	495	MetLeuThrTripleuSerValGlyGlyThrPheLeuSerTyThrMetValValSerAlaPheThr	514
Db	960	ATGCTCACCTGCAGTGTGGAGGCACCTTTCTCTCGTACATGTTGTGAGCGCCTTACC	1019

Qy	515	lleValPheValValLeuTripleuValProGluThrIysGly***AsnSerArgGlyAspThr	534
Db	1020	ATCGTTTTTGTTCCTTTGGTCCCGGAGACGAGGGAG-AACTCTAGAGGAGATACA	1078
Qy	535	llePheValSerLeuSerileGlnArgGlnLeuGlnTripleuProGluCysLeuSer	553
Db	1079	ATTTTCGTTTCGTGAGCAITTCAGCTCAGTCAATGGTTGCCCGAGTGTATATCT	1135

RESULT 8

US-10-424-599-132422

; Sequence 132422, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 132422

; LENGTH: 2202

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(2202)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MFT3847\_90S94C.1

US-10-424-599-132422

Alignment Scores: 2.23e-161 Length: 2202

Pred. No.: 1718.00 Matches: 334

Score: 83.67% Conservative: 76

Percent Similarity: 68.16% Mismatches: 76

Best Local Similarity: 61.82% Indels: 5

Query Match: 17 Gaps: 2

DB:

US-10-051-909-36 (1-553) x US-10-424-599-132422 (1-2202)

Qy	52	GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71
Db	255	GAAGAGAGTGGCGATCGGAGGAT-----CTTCAGAACCGTTCTCCACACGGAGT	308
Qy	72	TriTyArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
Db	309	TGTTACAAAGATGGTTCCAGGCGATCCAGCATCATG-----GGATCCTCCACTCATGT	362
Qy	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuileValAlaLeuGlyProile	111
Db	363	ATCCGCGACGGCGCCCTTCCGTCCTCTTCGGTCCTCATCTCGCTTGGTCCCAT	422
Qy	112	GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu	131
Db	423	CAATTGGCTTTCAGTGTGGTATTTCTTCCAAACCAAGGGCTATAGTTCGCGATCTA	482
Qy	132	AsnLeuSerileSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal	151
Db	483	AACCTCTATTTTCGGAGTTTCTTCTTGGATCTTTGCTTAATGTGGAGCAATGGTG	542
Qy	152	GlyAlaileAspGlyMetAlaGluTyThrileGlyArgGlyGlySerLeuMetile	171
Db	543	GGAGCTATAGTGTGGTTCAGTAGTGAATACATCGGCGCAAGGGTCAATGATGAT	602
Qy	172	AlaAlaileProAsnileGlyTripleuAlaileSerPheAlaLysAspAlaSerPhe	191
Db	603	GCTCGCATCCCAATATATATAGSGTGGCTTGCTATTTCTTTTGCCAAAGATTCCTCGTT	662

```
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 663 TTGTATATGGGAGGTGTTGGGAAGGTTTGGCGTATGTTATCTTATGTGGTGCCT 722
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 723 GTTTATATAGTCAGATTGACCTCAAACTTGAGAGGTGGCTTGGGTCACTGAAACCAG 782
QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrIleLeuGlyMetPheIleProTyrArg 251
Db 783 CTCTCTGTTACAATTGGCATTATGCTGCTTATCTGTGGGTCTTTTGTCAACTGGAGA 842
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271
Db 843 GTCTCTGCAATTCTAGGAATTTGGCCCTGTACAGTATTAATACCTGGATTATTTTCATA 902
QY 272 ProGluSerProArgTyrPheLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db 903 CCTGATCCCCAGATGGTTGGCCAGATGGGATGATAGATGAGTTTGAGACTTCTTTG 962
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db 963 CAAGTGTACAGGATTTGACACTGATATATCTGTTGAAGTACATGAAATTAAGAGATCT 1022
QY 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
Db 1023 GTGGCTTCAACGGGNAAGAGCTGCAATCCGATTTTGCAGATCTCAAGAGGNAAGATAT 1082
QY 332 ArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsn 351
Db 1083 TGGTCCCGTAAATGGTGTGATGTGATTAATCTTCTTCCAGCAATATCTGTGATCAAT 1142
QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
Db 1143 GGAATTTGTTCTATTCAACTACCACTTTTGCAATGCAGAAATTTCAATCCAGGAAGCT 1202
QY 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrProLeuLeu 391
Db 1203 GCTACAGTTGGACTTGGAGCGTTTCAGGTATAGCAACTGGAATTTCCACATGTTGGTG 1262
QY 392 AspArgAlaGlyArgGlyLeuLeuIleSerThrSerGlyMetThrLeuCysLeu 411
Db 1263 GACAAAGTGGCGGAGGTGCTTCTTAATAATATCTCTCAATGATGACATGTTAGCCTT 1322
QY 412 LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
Db 1323 CTCAATTTGTTCTATAGCATTTTATCTGGAGGGGTTGTATCAGAGGATTCACATTTATTC 1382
QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
Db 1383 AGCATTTTGGGAATAGTTTCTATTGTTGGACTCGTGCTATGATGGGTCTCTCTA 1442
QY 452 GlyMetGlyAlaIleProTyrPheLeuMetSerGluIleLeuProValSerIleLysSer 471
Db 1443 GGTCTGGGACCCATCCCTGGCTTATAATGCTGAGATACCTCCAGTGAATATAAAGGGC 1502
QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTyrPheLeuThrSerPheAlaIleThrMetThr 491
Db 1503 CTTGCTGGCAGCATAGCAATAGGGAATTTGGCTGATTTCTGTTGGGGATCAGATGACT 1562
QY 492 ThrAsnLeuMetLeuThrTyrProValGlyIleThrPheLeuSerTyrMetValValSer 511
Db 1563 GCTAACTTGTGTTTGAATTTGGAGCAGTGGAGGACATTTACAATCTACACAGTCGTAGCT 1622
QY 512 AlaPheThrIleValPheValValLeuTyrProGluThrLysGly***AsnSerArg 531
Db 1623 GCCTTTTACTATAGCTTTTATAGCAATGTGGGTTCCTGAGACCAAGGGAAG-AAACATTGGA 1681
QY 532 GlyAspThrIlePheValSerLeuSerIle 541
Db 1682 AGAAATTCAGTTTCTTTCGATAGATATA 1711
```

## RESULT 9

```
US-10-437-963-63465
; Sequence 63465, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21 (53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 63465
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64704C.1
US-10-437-963-63465

Alignment Scores:
Pred. No.: 2,896-127 Length: 1304
Score: 1374.00 Matches: 293
Percent Similarity: 68.89% Conservative: 48
Best Local Similarity: 59.19% Mismatches: 58
Query Match: 49.44% Indels: 97
DB: 18 Gaps: 6

US-10-051-909-36 (1-553) x US-10-437-963-63465 (1-1304)
```

```
QY 51 GlyGluGluSerGlySerAspHisAspGly-----ValLeuArgArg 64
Db 72 GCGAGGAGAGCGCGCGGAGGATGGGGCGGCGCGCTCGCGCTCGGACCTCGCGAAG 131
QY 65 ProLeuLeuAsnThrGlySerTyrTyrArgMetSerSerArgGlnSerSerPheAlaPro 84
Db 132 CGTTTCTCCACACGGGGAGCTGTGTACAGATGTCTGCGCGGCGCGCGCGGGATG 191
QY 85 GlyThr-----SerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeu 101
Db 192 GGGTCGCGCTCGGCTCTCCGCTACTCCCTCCGCGACTCTCTCCGCTCTCCGCGCTCTC 251
QY 102 CysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSer 121
Db 252 TGCACCTCATCGTCGCCCTCGGCCCATCCAGTTCGGCTTCACCTGCGGCTTCTCCTCG 311
QY 122 ProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPhe 141
Db 312 CCACCCAGAGCGCCATCATCTCCGACCTCGGCTTCACCTCTCCGAGTTCTCGTGTTTC 371
QY 142 GlySerLeuSerAsnValGlyMetValGlyValAlaIleAlaSerGlyGlnMetAlaGlu 161
Db 372 GGGTCGCTGTGNAAGCTCGGGGCAATGGTGGGGCCATGGCCAGCGGCCAGATCGCGAG 431
QY 162 TyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTyrLeu 181
Db 432 TACATCGCGCGCAAGGGTCTCTCATGATCGCGCAATTCCTAACATTATTGGGTGGCTC 491
QY 182 AlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyPhe 201
Db 492 GCAATATCTTTTGCAAGGACTCA----- 515
QY 202 GlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsn 221
Db 516 -----TGAGTCCAGTTTATATAGCAGAAATTTGCTCCACAACA 554
```



Db 681 GAGATATCTCACAGAACATGAGAGGGGCTTGGCTCTGTGAACCAAGTTGTCTGAACC 740  
Qy 236 PheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaVal 255  
Db 741 CTGTGATATCATGTTGGCTATTTGCTCGGCTTGTGTTGTTCTTGGAGGCTTCTTGCAATA 800  
Qy 256 IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerPro 275  
Db 801 ATAGGAACCTTGGCCCTGCATAGTGTTCATACCTGGCCCTTTTCTTCATTCCGGAATCTCCA 860  
Qy 276 ArgTyrLeuAlaValMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295  
Db 861 AGATGGCTGCGCAGGATGAATATGATGGATGATTCGAGACTTCTTACCAAGTTCTGAGA 920  
Qy 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleValArgAlaValAlaSerSer 315  
Db 921 GGATTCGATGCTGACATCACTGCAGAAATTCGAATGATATAAGAGAGCAGTAATGTCAGCA 980  
Qy 316 SerIleArgThrThrIleSerPheGlnGluLeuAsnGlnIleValTyr----- 331  
Db 981 AACAAAGGGGCTCAATTCGTTTCCAGAGTTGAACTTGAACCAAAAAAAAAAAAAAGCGGC 1040  
Qy 332 ArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsn 351  
Db 1041 CGTACGCCCCCTAATCTAGGAATGGGCTACTTGTGTGAGCAGCTGAGTGGGATAAAC 1100  
Qy 352 GlyValLeuPheTyrAlaSerSerIlePheValAlaAlaGlyValThrAsnSerAspLeu 371  
Db 1101 GGTATAATCTTTTATGAGTAGTACGACTT-AAAGCTGCAGGCTCAAGACAGTAACCTG 1159  
Qy 372 AlaThr 373  
Db 1160 GACACA 1165  
RESULT 11  
US-10-425-115-47715  
; Sequence 47715, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 47715  
; LENGTH: 1105  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_143523C.1  
US-10-425-115-47715  
Alignment Scores:  
Pred. No.: 3,26e-121 Length: 1105  
Score: 1312.50 Matches: 271  
Percent Similarity: 92.31% Conservative: 5  
Best Local Similarity: 90.64% Mismatches: 10  
Query Match: 47.23% Indels: 16  
Gaps: 1  
DB: 1  
US-10-051-909-36 (1-553) x US-10-425-115-47715 (1-1105)  
Qy 1 ProSerSerSerSerPheArgProAlaGlyLysLysValLysValLysAsnGlnGly 20  
Db 212 CCTTCCTCTCTGCTCTCTTCAGGCCGCGGAGAGAGAGA--AAATATCAGGC 269  
Qy 21 LeuArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGly 40

Db 270 TTGCGGCGAGAGGCTGTGCGCGCCGACCGCGGAGGAACTTCTGTACGCCGCTCATGGT 329  
Qy 41 GlyIleSerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAspGly 60  
Db 330 GCGCGCAGCAACAGAGCGCGCGCGCGCGGAGAGAGCGGAGCGACACGACGCT 389  
Qy 61 ValLeuArgArgProLeuLeuAsnThrGlySerTyrArgMetSerSerArgGlnSer 80  
Db 390 GTGCTCGGAGGCGCTGTCTCAACACGCGGAGCTGTGTACCGGATGAGCTCGCGCAGTCC 449  
Qy 81 SerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPhe 100  
Db 450 AGCTTTGCGCGGAGGACTCTCTCCAG-GCGCTCTGCGGAGTCCACAGTCTCCGCTTC 508  
Qy 101 LeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSer 120  
Db 509 CTCTGCAGCGCTCATCGCTCGCGCTCGGCCCATCCAGTTGCGCTTCCACGCGCTTCTCC 568  
Qy 121 SerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAla 140  
Db 569 TCCCGACCCAGGACGCGCATGGTTGCGGACCTCAACCTCTCTATCTCCGAGTTCTCGCGC 628  
Qy 141 PheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAla 160  
Db 629 TTGCGATCGCTGTCCAACTGCGCGCGCATGGTGGGCGCATCCGCGGAGAGATGGCC 688  
Qy 161 GluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTyr 180  
Db 689 GAGTACATGGCGGTAAGGGTCTGTGATGATGCTGCAATCCCAATATCATCGGTTGG 748  
Qy 181 LeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGly 200  
Db 749 CTTCGATCTCCTTTCGAAAGATGCTCATTTCTGTATATGGGAGCGCTTGTCTTGAAGA 808  
Qy 201 PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln 220  
Db 809 TTGCGGTGTCGGTATCATTTCTCTATGTTGGTACCGGTATACATAGCAGAGATATCTCCAA 869  
Qy 221 AsnMetArgGlyAlaLeuGlySerValAsnGlnLeu----- 232  
Db 869 AACATGAGAGGGGCTCTTGGCTCTGTGNACCAGGT-TTGCTATACTTGCAGCATCTCA 927  
Qy 233 -----SerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247  
Db 928 TCTATTGTGAGTTATCTGTAACTTGGTATCATGTTTGCCTATTGCTCGGCTTGT 987  
Qy 248 IleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267  
Db 988 GTTCCTTGGAGGCTTCTTGAGTAATAGGAACCTTGCCTGTCATAGTGTGATACCTGGC 1047  
Qy 268 LeuPhePheIleProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAsp 286  
Db 1048 CTTTCTTCTTCCGGAATCTCAAGATGGCTGCGCAAGATGAATATGATGGATGAT 1104  
RESULT 12  
US-10-425-114-28778  
; Sequence 28778, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven B.  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 28778

```

; LENGTH: 957
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4764-010-D11_FLI
US-10-425-114-28778

```

```

Alignment Scores:
Pred. No.: 3,41e-112 Length: 957
Score: 1221.50 Matches: 249
Percent Similarity: 85.27% Conservative: 23
Best Local Similarity: 78.06% Mismatches: 33
Query Match: 43.95% Indels: 14
DB: 17 Gaps: 6

```

US-10-051-909-36 (1-553) x US-10-425-114-28778 (1-957)

```

QY 23 ArgGluAlaValProGlyArgPro-----AlaSerGluLeuArgThrArgValMetGly 40
DB 2 CGGGAACGGTCCACCCACCGGAGAAACGAAACAAACCAATCTCGTGGAGCTG 61
QY 41 GlyGlySerAsn-ArgGlyGly-----AlaGlyAlaGlyGlu---G1 53
DB 62 GCGGCGCGGCGCGCGCGCCATGCGGAGGAGGAGGAGCTCGGCGCGCGAGTAGCA 121
QY 53 uSerGlySerAsnHisAspGlyValLeuArgArgProLeuLeu---AenThrGlySerTr 72
DB 122 GAGCGGCGGAGGAGGAGGCGGCTGAGAGCGGCTGAGAGCGGCTGCTGCCCAACAGCGGAGCTG 181
QY 72 pTyrArg-----MetSerSerArgGlnSerSerPhe---AlaProGlyThrSerSerMe 89
DB 182 GTACCGGATGGGATGGGTCGCGCGAGTCCAGCTCCACCGCGCGCGCACCTCTCCAT 241
QY 89 tAlaValLeuArgGluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuG1 109
DB 242 GCGCGCTCGTGGAGTCCACGCTCCCGCTTCCTGCGAGCTCATGTCGCGCTCGG 301
QY 109 YProGlnGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetVal 129
DB 302 CCCCATCCAGTTCGGCTTCCCGCGCGCTTACTCTCCCGCGAGCGAGCGCATCCCG 361
QY 129 gAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyG1 149
DB 362 GGACCTCAACCTCTCCATCTCCGAGTTCCTGGGTTCGGCTCGCTGTCACAGTGGGCGC 421
QY 149 YMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerIe 169
DB 422 CATGTCGGGCGGATGCCAGCGCGCCAGATGCCGAGTACGTTGGCGCAAGAGCTCT 481
QY 169 uMetIleAlaIleProAsnIleGlyTrpLeuAlaIleSerPheAlaLysAspAl 189
DB 482 GATGATTGAGCAATTCGAACGCTCAITGGTTGGCTTGGCATCTCTTTGCAAGAGCTC 541
QY 189 aSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleSerTyrTh 209
DB 542 TTCAATTTCTATATATGGGAGCTTGTGAAGGATTCGGTTCGGTATCATTTCTATGT 601
QY 209 rValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVa 229
DB 602 GGTACCGGTATACATAGCAGATATCTCCACAAAACATGAGGGGGCTCTTGGCTCTGT 661
QY 229 lAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePr 249
DB 662 GAACCACTTATCTGTAACTTCCATGATGTTGCTTACTTGTGCTTGTGCTTGTGTTCC 721
QY 249 oTrpArgLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPh 269
DB 722 TTGGAGGCTTCTTGCAGTATAGGAACCTTGGCCCTGCATAGTGTGTGATACCTGGCTTT 781
QY 269 ePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluTh 289
DB 782 CTTCTTCCGGATCTCCAGATGGCTGGCAAGATGATATGATGATGATGATGATGATGATG 841

```

```

QY 289 rSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLy 309
DB 842 TTCTCTACAAGTTCTGAGAGGATTCGATGCTGATCCTCCTGCGAATGATGATATAA 901
QY 309 sArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsn 327
DB 902 GAGAGCAGTAATGTCAGCAACAAAGGGCTACAAATTCGTTTCCAGAGTTGAAC 956

RESULT 13
US-10-424-599-131066
; Sequence 131066, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: Ia Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Xongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131066
; LENGTH: 1283
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8935C.1
US-10-424-599-131066

Alignment Scores:
Pred. No.: 2,39e-108 Length: 1283
Score: 1185.00 Matches: 223
Percent Similarity: 81.84% Conservative: 61
Best Local Similarity: 64.27% Mismatches: 51
Query Match: 42.64% Indels: 12
DB: 17 Gaps: 2

US-10-051-909-36 (1-553) x US-10-424-599-131066 (1-1283)
QY 201 PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln 220
DB 2 TTGGGGTATAGGAATATATCTTACACGGTCTCTGTACATAGCTGAGATATACCTCCA 61
QY 221 AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu 240
DB 62 AACTTACGGGGGGGTTTGGTTTTCAGTTAACCAAGCTCTCTGCACCATTTGGAAATATGCTG 121
QY 241 AlaTyrLeuLeuGlyMetPheIleProThrArgLeuLeuAlaValIleGlyAlaLeuPro 260
DB 122 GCATATCTCTGGGATTTTGTGATGGAGAAATCTTGCATTTCCAGTCCCTAGATGGCTGCAAAA 181
QY 261 CysThrMetLeuIleProGlyLeuPheIleProGluSerProArgTrpLeuAlaLys 280
DB 182 TGTAACAATATGATACCTGGCTATTTTTCATTTCCAGTCCCTAGATGGCTGCAAAA 241
QY 281 MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAsp 300
DB 242 AUGGAATGACAGAGAATTTGAACTCTCTTGAAGTCTCCGAGGCTTTGAGACTGAT 301
QY 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr 320
DB 302 ATTCTGTGTGAGTGAATGAAATTAAGAGGCTGTGTGCTTCAACAACAGAGAACTACA 361
QY 321 IleSerPheGlnGluLeuAsnGlnLysTyrArgThrProLeuLeuGlyIleGly 340
DB 362 GTTCGATTTGACAGCTCAACAAAGAGATATGCTTCCCTTCCCTTATGATGGAATGGA 421
QY 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerIle 360
DB 422 CTGCTTATTTTTCACAGCTTTCTGGAATTAATGGTGTCTTTTATTTTCCAGTACATC 481

```





1018 AGCTACCTTTGGAGTTGGCTTCATTCAGGTCAATGCTACTGCTTTAAACACACATGGTTG 1075

RESULT 15  
US-10-424-599-132426  
; Sequence 132426, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 132426  
; LENGTH: 2105  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(2105)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_90588C.1  
US-10-424-599-132426

Alignment Scores:  
Pred. No.: 1,79e-103 Length: 2105  
Score: 1139.50 Matches: 228  
Percent Similarity: 70.55% Conservative: 69  
Best Local Similarity: 54.16% Mismatches: 58  
Query Match: 41.00% Indels: 67  
DB: 17 Gaps: 2

US-10-051-909-36 (1-553) x US-10-424-599-132426 (1-2105)

QY 187 LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle 205  
Db 320 CAGGATTCCTCATTTTGTATATAGGGAGCCCTGCTGGAAGGTTTGGCGTCGGGATTAGC 379  
QY 207 SerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeu 226  
Db 380 TCCTATTAGTGCTGTTTATATAGCTGAGATTGACCTCAAACTTCGAGAGGTGGCCTT 439  
QY 227 GlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuGlyMet 246  
Db 440 GGATCAGTGAACGAGCTCTATATCTATTGTCATTTGTCATTTGCTTATCTTTTGGCCTT 499  
QY 247 PheIleProTyrPheArgLeuAlaValIleGlyAlaLeuProCysThrMetLeuIlePro 266  
Db 500 TTTGTCAACTGGAGAGTGGTTCGAATTCAGGAATTTTGGCCCTGTACAGATTATTAACCT 559  
QY 267 GlyLeuPhePheIleProGluSerProArgTyrLeu 278  
Db 560 GGATTATTTTTCATACCTGAATCCCCCAGATGGTTGGTATGGATATGTTATGCATTATAT 619  
QY 278 278  
Db 620 TGATCAAGTTTTCATCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 679  
QY 278 278  
Db 680 GTAACTCTTCTACTCTTGGTAATTATCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 739  
QY 279 279  
Db 740 TTTTCTGTTGTTGGGAATCGGGAAGTGTAGGCTAGAGTGGGATGACAGATGAGTT 799  
QY 287 sGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAs 307  
Db 287 287

Db 800 TGAGACTTCTTTGCAAGTGTTCAGAGATTGTGACACAGATATATCTGTTGAGATATATGA 859  
QY 307 pileLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAs 327  
Db 860 AATTAAGAGATCTGTGGCTTCAACGGGAAAAAGAGCTACAAATCCGATTTTCAGATCTCAA 919  
QY 327 nGlnLysLysTyrArgThrProLeuLeu 341  
Db 920 GAGGAAAAGATATTGGTTCCCTTAATGGTATATATGACACTGTGATTCGTTATGGATT 979  
QY 341 uLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePh 361  
Db 980 ACTTGTCTTTCAGCAGTTATCTGTATCAATGGAGTTTGTCTTATCAACTACCATCTT 1039  
QY 361 eLysAlaIleGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnVa 381  
Db 1040 TGCAAAATCGAGGAATTTGTCAGGAAAGTGTCTACTGTGGACTTGTGTCGGCTTCAGGT 1099  
QY 381 lLeuAlaThrGlyValThrThrTyrLeuLeuAspArgAlaGlyArgIleLeuIle 401  
Db 1100 CATAGCAACTGGAATTTCCACGTGGTGGTGGACAAAGTGTGCGAGGCTGCTTCTAT 1159  
QY 401 eIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLy 421  
Db 1160 GATATCCTCATCTGTATATGACTGTAGCCTTCTCATTTGTTCAATAGCATTTTATCTGGA 1219  
QY 421 sAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGl 441  
Db 1220 GGGGTTGTATCTGAGGATTTCATCTATTCAGCATGTGGGAATAGTTTCTGTGTGTGG 1279  
QY 441 YIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMe 461  
Db 1280 ACTTGTGGTTATGTCATTTGGGTTCTCTCTGTCGGACCCATCCCTTGGCTTATAAT 1339  
QY 461 tSerGluIleLeuProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAs 481  
Db 1340 GTCTGAGATACTTCCAGTGAAATTAAGGCTTGTCTGGCAGCATAGCGCAATGGGAAA 1399  
QY 481 nTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGl 501  
Db 1400 TTGGCTGATTTTCGTGGGTGATCAGCATGACCGTAACCTTACTTTTGAATTTGGAACAGTGG 1459  
QY 501 YGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTr 521  
Db 1460 AGGGACATTTACAATCTACACAGTGGTAGTGCCTTTACTATTCCTTTTATAGCATATG 1519  
QY 521 pValProGluThrLysGly\*\*\*AsnSerArgGlyAspThrIlePheValSerLeuSerIl 541  
Db 1520 GGTTCCTGAGACCAAGGAAG-AACATTTGGAAGAAATTCAGTTTTCCTTCAGATAGATGT 1578  
QY 541 e 541  
Db 1579 T 1579

Search completed: April 14, 2005, 01:49:24  
Job time : 731.592 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 12:44:19 ; Search time 3817.87 Seconds

(without alignments)  
5513.431 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779

Sequence: 1 PSSSSFRPAGKKKKKNG.....TIFVLSIQRLQWLPECL 553

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp  
-Q/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074037\_14060/app\_query.fasta\_1.1678  
-DB=EST -QMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOGPCL=0 -LOOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFTW=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN 1.1 6628 @runat\_13042005\_074037\_14060 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733.5	62.4	1837	3	CNSOACUK
2	1219	43.9	854	7	CN125032
3	1102	39.7	868	6	CB675064
4	1065	38.3	682	6	CA191028
5	1021	36.7	635	6	CA143053
6	1004.5	36.1	764	6	CB648359
7	984.5	35.4	1155	7	CK211005
8	984	35.4	674	1	AV939950
9	976	35.1	714	6	CA262285

10	975	35.1	681	1	AV913420
11	962.5	34.6	718	7	CR291514
12	962.5	34.6	762	1	AJ794429
13	932	33.5	701	6	CD871121
14	930.5	33.5	1675	3	CNSOAHU
15	925	33.3	769	4	BI933646
16	919.5	33.1	1737	3	CNSO92VC
17	914	32.9	772	6	CB683315
18	907.5	32.7	1666	3	CNSOAHV
19	907	32.6	666	5	BQ766951
20	904.5	32.5	627	6	CA180355
21	904	32.5	652	6	CA201877
22	902	32.5	716	7	CF451024
23	901.5	32.4	748	6	CB347677
24	896.5	32.3	1601	3	CNSOAG20
25	896	32.2	685	5	BQ862009
26	893	32.1	566	4	BM325827
27	890	32.0	791	6	CB892803
28	887.5	31.9	746	6	CD483180
29	882	31.7	594	6	CD203945
30	882	31.7	772	7	CF835377
31	870.5	31.3	744	1	AJ796408
32	869	31.3	641	6	CA246696
33	864.5	31.1	927	7	CK153511
34	859	30.9	548	6	CA143695
35	850	30.6	627	4	BJ249193
36	847	30.5	663	7	CN906216
37	845	30.4	670	5	BQ855360
38	838	30.2	595	2	BE599181
39	830.5	29.9	651	4	BG522368
40	828	29.8	667	6	CD893209
41	827	29.8	626	7	CN904481
42	825	29.7	656	5	BQ913239
43	822	29.6	1054	3	AY111571
44	821	29.5	686	7	COL16723
45	820.5	29.5	714	6	CA262252

#### ALIGNMENTS

RESULT 1  
CNSOACUK

LOCUS  
DEFINITION

CNSOACUK 1837 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTSL242B10 of Adult vegetative tissue of strain col-0 of  
Arabidopsis thaliana (thale cress).

ACCESSION

VERSION  
BX815012.1 GI:42472886

KEYWORDS

SOURCE  
HTC; GSUT CDNA.

ORGANISM

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schacher, V.,  
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

TITLE

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome

JOURNAL

REFERENCE

Unpublished  
2 (bases 1 to 1837)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT

- Web : www.genoscope.cns.fr  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

AV913420	AV913420
CR291514	CR291514
AJ794429	AJ794429
CD871121	AZ02.117H
BI933646	Arabidops
EST553535	
EX8132192	Arabidops
CB683315	OSJNE111
EX819760	Arabidops
BQ766951	EBR08 SQ
CA180355	SCCST300
CA201877	SCRPL102
CF451024	EST687369
CB347677	CB22SG000
EX222441	Arabidops
BQ862009	QGC1F12.Y
BM325827	PIC1.53.C
CB892803	EST645595
CD483180	atr01-31m
CD203945	H51.2_E01
CF835377	UCRCS03.0
AJ796408	AJ796408
CA246696	SGSGFL5C0
CK153511	FGAS03213
CA143695	SCRUT200
BJ249193	BJ249193
CN906216	010929ABD
BQ855360	QGB25P02.
BE599181	PI1.86.D0
BG522368	20-64.Ste
CD893209	G118.123C
CN904481	010831ABD
BQ913239	QHA6L20.Y
AY111571	Zea mays
COL16723	GR EB019
CA262252	SCBQLB201





```

/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSUNBe"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

```

## ORIGIN

Alignment Scores: 2,01e-103 Length: 868  
 Pred. No.: 1102.00 Matches: 208  
 Score: 87.89% Conservative: 46  
 Best Local Similarity: 71.97% Mismatches: 35  
 Query Match: 39.65% Indels: 0  
 DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CB675064 (1-868)

```

QY 206 IleserTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
DB 1 ATATCTATGTGTGTCAGTTATATAGCAGAAATGCTCCACAAACAATGAGAGAGCT 60
QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuGly 245
DB 61 CTGGTTCAGTCATCAGCTTTCGTACTATTGTCATATGCTTGTCTACTTGTAGGC 120
QY 246 MetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
DB 121 ATGTTTGTCCCTGGAGAAATCTCTCCGTCCTCGGTATTTTACCAATGTTCAATTCGATA 180
QY 266 ProGlyLeuPhePheIleProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGlu 285
DB 181 CTGGATTTGTTTATATCTCCTGAATCACCAGAGTGGCTGGCAAAATGGGAAGATGAAG 240
QY 286 AspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluVal 305
DB 241 GATTTTGAATCTCTCGCTGCAAGTATTGCGAGGATTTGAACAGACATCGCTGTAGAAGTA 300
QY 306 AsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGlu 325
DB 301 AATGAAATAAGAGAACAGTTCATCATCACGGAGGAGGACGACAAATACGATTTGCAGAT 360
QY 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGln 345
DB 361 ATCAAGCAGAGAGATATAGTACCTCTAATGATAGGAATGGTCTCTTGTACTGCAG 420
QY 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
DB 421 CAGTTAAGTGTGTGAATGCGATTCATTTTATGCTGCTAGTATCTTCAAGCCGCCGT 480
QY 366 ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly 385
DB 481 CTTACGAATAGTAAATACACACATTTGTCCTGGGGGTTGTCAGGTGGTGTCTACTGGA 540
QY 386 ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSer 405
DB 541 GTGACAACTTGGTTGATCAGCAAGGTGGTCAGCACTCTCTCTATTTATCTTACTACA 600
QY 406 GlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSer 425
DB 601 GCGATCACCATTACTCTTGTGTGTGTCTGTGTCTATTTTGTGTAAGCAGCAACATRACT 660
QY 426 GluAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe 445
DB 661 AATGGTCTCATTTATATCTCTGTAATAGTATGCTTGTGCTGGTGGGCTTGTGGCAATT 720
QY 446 ValIleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeu 465
DB 721 GTGATTCATTTTCTCTTGGTTGGAGCCATACCATGATCATATAATGCTGAGATTCCT 780
QY 466 ProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer 485
DB 781 CCTGTAATATCAAGAGCCCTTCTGGAAGCGCTTGCTGCAACCTTGCAACCTGGTGTAGCGCT 840

```

```

QY 486 PheAlaIleThrMetThrThrAsnLeu 494
DB 841 TGGCTCATTCAGTACGATGACAGCAAGCTTG 867

RESULT 4
LOCUS CA191028
DEFINITION 5', mRNA sequence.
ACCESSION CA191028
VERSION CA191028.1 GI:35135776
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
            Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE 1 (bases 1 to 682)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccnetcenter.fcav.unesp.br
            Plate: C03 row: A column: 08
            Seq primer: 17 Promoter Primer.
            Location/Qualifiers
                1..682
                /organism="Saccharum officinarum"
                /mol_type="mRNA"
                /db_xref="taxon:4547"
                /clone="SCCCT2C03A08"
                /lab_host="DH10B"
                /clone_lib="RT2"
                /note="Organ: Root tips (0.3cm-long) from adult plants;
                Vector: pSport1; Site 1: SalI; Site 2: NotI; An
                unidirectional cDNA library generated from [Root
                tips (0.3cm-long) from adult plants]. cDNA was prepared
                from polyA+ mRNA using SuperScript Plasmid System Kit
                (Invitrogen). The double-strand cDNAs were fractionated
                in a sepharose CL-2B 40cm-columns and fragments sizing
                between 0.8 and 1.5 Kb were directionally cloned into the
                vector. Details of each source of RNA and library
                construction can be obtained at
                http://sucest.lad.ic.unicamp.br/public"

```

## ORIGIN

Alignment Scores: 9,58e-100 Length: 682  
 Pred. No.: 1065.00 Matches: 218  
 Score: 97.36% Conservative: 3  
 Best Local Similarity: 96.04% Mismatches: 6  
 Query Match: 38.32% Indels: 1  
 DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CA191028 (1-682)

```

QY 256 IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGluSerPro 275
DB 2 ATCGGAGCCTTCCTCGTCAGACGTGTGTGATTCCTGTGACTATTCTTCATTCAGAAATCTCCC 61
QY 276 ArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295
DB 62 AGATGCTGGCAAGATGAATTTGATGGAAGATTCGAGACGTCCTCCTCAAGTGTCTGAGG 121

```

QY 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer 315  
 DB 122 GGGTTTGGAGACTGC-ATCAGCAGAGAAGTGAATCATATAAGAGGCGAGTGACATCATCA 180  
 QY 316 SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysIleValArgThrProLeu 335  
 DB 181 AGTAAGAGGACTCAATCAGTTTTCAGAAATTAAACCAAGAAATACCCACGCGCTA 240  
 QY 336 LeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355  
 DB 241 CTTCTAGGAATTGGCCTACTTGTACTGCAAAATCTAAGTGAATCAACGGTGTACTGTTT 300  
 QY 356 TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer 375  
 DB 301 TATGCAAGTAGCATCTTCAAAGCTGCAGGTGTACAAACAGCAGACTTGGCCACCTGTTC 360  
 QY 376 LeuGlyValIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGly 395  
 DB 361 CTTGGAGCTATCAGGTCTCTGTACTGAGATTACGACATGGTTGTAGACAGCTGGA 420  
 QY 396 ArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415  
 DB 421 CGACGATGCTTCTCATTTATTTCTACCTCTGGCATGACTCTATGCCTTCTTGGCGTTCT 480  
 QY 416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435  
 DB 481 GTTGTATTTTCTCAAGGATAAGATTTCACAGGATTTCTAACTCGTACTACATCTTAAC 540  
 QY 436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455  
 DB 541 ATGATCTCTTGTGTGCTATCGTGGCTTTTGTCAATACCTTCTCGTTTGGTATGGGTGCC 600  
 QY 456 IleProTrpLeuMetSerGluIleLeuLeuProValSerIleLysSerLeuGlyLys 475  
 DB 601 ATTCCATGGCTCATGATGCTGAGATCTTCCAGTTAGCATCAAGAGTCTCGCGGAGC 660  
 QY 476 IleAlaThrLeuAlaAsnTrp 482  
 DB 661 ATCGGACACTGGCCAAATGG 681

RESULT 5  
 LOCUS CA143053  
 DEFINITION SCQRT2032A08.g RT2 Saccharum officinarum cDNA clone SCQRT2032A08 5', mRNA sequence.  
 ACCESSION CA143053  
 VERSION CA143053.1 GI:35038703  
 KEYWORDS EST.  
 SOURCE Saccharum officinarum  
 ORGANISM Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.  
 REFERENCE 1 (bases 1 to 635)  
 AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
 TITLE The libraries that made SUCEST  
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 COMMENT Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@unicamp.br  
 Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bccc.unesp.br>  
 Plate: 032 row: A column: 08  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1..635

FEATURES  
 source

/organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

/clone="SCQRT2032A08"

/lab\_host="DH10B"

/clone\_lib="RT2"

/note="Organ: Root tips(0.3cm-long) from adult plants;

vector: pSPori; Site\_1: SalI; Site\_2: NotI; An

unidirectional cDNA library generated from [Root

tips(0.3cm-long) from adult plants]. cDNA was prepared

from polyA+ mRNA using SuperScript Plasmid System kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

<http://sucest.lad.ic.unicamp.br/public>"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 3.18e-95 Length: 635  
 Score: 1021.00 Matches: 204  
 Percent Similarity: 98.10% Conservative: 3  
 Best Local Similarity: 96.68% Mismatches: 4  
 Query Match: 36.74% Indels: 0  
 DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CA143053 (1-635)

QY 256 IleGlyValAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerPro 275

DB 2 ATCGGAGCCTTGGCCCTGCACAGTTGATTCCTGGACTATTCTTCATTCCAGAACTCTCCC 61

QY 276 ArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295

DB 62 AGATGGCTGGCAAGATGATTTGTATGGAAGATTGGAGACGTCCTTACAAAGTGTGAGG 121

QY 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer 315

DB 122 GGGTTTGGAGACTGACATCAGCAGAGAAGTGAATGATATAAGAGGCGAGTGACATCATCA 181

QY 316 SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysIleValArgThrProLeu 335

DB 182 AGTAAAGAGACTACATATCAGTTTCAAGAAATTAACCAAGAAATACCGCAGCGCGCTA 241

QY 336 LeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355

DB 242 CTTCTAGGAATTGGCCTACTTGTACTGCAAAATCTAAGTGAATCAACGGTGTACTGTTT 301

QY 356 TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer 375

DB 302 TATGCAAGTAGCATCTTCAAAGCTGCAGGTGTACAAACAGGACCTTGGCCACCTGTTC 361

QY 376 LeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGly 395

DB 362 CTTGGAGCTATCCAGGTCTCTGTACTAGGATTACGACATGGTTGTATAGACAGCTGGA 421

QY 396 ArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415

DB 422 CGACGATGCTTCTCATTTATTTCTACCTCTGGCATGACTCTATGCCTTCTTGGCGTTCT 481

QY 416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435

DB 482 GTTGTATTTTCTCAAGGATAAGATTTTCACAGGATTTCTAACTCGTACTACATCTTAAC 541

QY 436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455

DB 542 ATGATCTCTCTGGTTGCTATCGTGGCTTTTGTATTACCTTCTCGTTTGGTATGGGTGCC 601

QY 456 IleProTrpLeuMetSerGluIleLeuPro 466

DB 602 ATTCCATGGCTCATGATGTCTGAGATCTCTTCCA 634

## RESULT 6

CB648359  
 LOCUS OSJNEB1110.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA  
 DEFINITION clone OSJNEB1110 5', mRNA sequence.  
 CB648359  
 ACCESSION CB648359.1 GI:29643352  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 764)  
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
 Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 Unpublished (2003)  
 CONTACT: Rod Wing  
 ARIZONA GENOMICS INSTITUTE  
 UNIVERSITY OF ARIZONA  
 BIOLOGICAL SCIENCES WEST, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 TEL: 520 626 3967  
 FAX: 520 621 9288  
 EMAIL: http://genome.arizona.edu

PCR PRIMERS  
 FORWARD: gta aac cga cgg cca gtc  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 11 row: L column: 10  
 Seq primer: gta aac cga cgg cca gtc.  
 Location/Qualifiers

1..764  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="OSJNEB1110"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSJNEB"  
 /notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:  
 XhoI; 24 hrs after inoculation with Rice Blast (Che  
 86061)"

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,21e-93 Length: 764  
 Score: 1004.50 Matches: 202  
 Percent Similarity: 87.40% Conservative: 13  
 Best Local Similarity: 82.11% Mismatches: 24  
 Query Match: 36.15% Indels: 7  
 DB: 6 Gaps: 2

US-10-051-909-36 (1-553) x CB648359 (1-764)

QY 28 GlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySerAsnArgGlyGly 47  
 Db 35 GGAGCCGAGCGAGCGAGCGGCGCGATGAACGCGCGGCGGAGCAGCGCGGAGGT 94  
 QY 48 AlaGlyAlaGlyGluGluSerGlySerAspHisAsp-----GlyValLeuArg 63  
 Db 95 CGCTGTCGCGCAGCAGAGCGGCGAGCGGCGGCGGCGGCGGCGGCTTGCGG 154  
 QY 64 ArgProLeuLeuAnthrGlySerTrpTyrArgMetSerSerArgGlnSerPheAla 83  
 Db 155 AAGCGCTGCTGAACAGCGGAGCTGGTACAGATGGATCGCGTCCAGCCTCGCGCC 214  
 QY 84 ProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThr 103  
 Db 215 -----TCTTCATGGCGCGCATCCGGAGTCCACAGCTCCCGCTTCTCTCGACG 265

QY 104 LeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThr 123  
 Db 266 CTCATCGTCGGCTCGGCCCATCATCAATTTCGATTCACCGCGCTTCTCTCACCACC 325  
 QY 124 GlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySer 143  
 Db 326 CAGGAGCCATCATCCGCGACCTCAAGCTCTCCATCTCCGAGTTCTCGGCTTCGGTTCG 385  
 QY 144 LeuSerAnValGlyGlyMetValGlyValAlaIleAlaSerGlyGlnMetAlaGluTyrIle 163  
 Db 386 CTGTCCACAGCTCGGCCCATCGTCGGAGCATGCCAGTGGCAGATGGCGGAGTACATT 445  
 QY 164 GlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIle 183  
 Db 446 GCCCGAAGGTCGTGATTAATTCGCGCGGTTCTTAACATCATTTGGTTGGCATC 505  
 QY 184 SerPheAlaLeuAspAlaSerPheLeuTyrMetGlyArgLeuGlyGlyPheGlyVal 203  
 Db 506 TCCTTTGCAAAAGACGCGTCATTTTATACATGGAGCGCTTGTGAAGTTTGGTGT 565  
 QY 204 GlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArg 223  
 Db 566 GGTGTATATCATATACGTCGCCATATACATAGCAGAGATATCTCATCAGACACAGA 625  
 QY 224 GlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeu 243  
 Db 626 CGAGCACTTGGCTCGGTGAACCCAGTTGTCCTTACCATTTGGTATCTTGTGGCCTATTG 685  
 QY 244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263  
 Db 686 CTAGGCACTTTGTTCTTGGAGGCTGCTTGCAGTGATAGGAAGCATCCCATGTACATTG 745  
 QY 264 LeuIleProGlyLeuPhe 269  
 Db 746 TTAATACCTGGTCTATTTC 763

RESULT 7  
 CK211005 1155 bp mRNA linear EST 08-DEC-2003  
 LOCUS FGAS022833 Triticum aestivum FGAS: Library 5 GATE 7 Triticum  
 DEFINITION aestivum cDNA, mRNA sequence.

CK211005

CK211005.1 GI:39573395

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Triticum aestivum (bread wheat)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.

1 (bases 1 to 1155)

Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,

Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,

Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,

Penniket,C., Roach,J.L. and Sarhan,F.

Functional Genomics of Abiotic Stress In Wheat and Canola Crops

Unpublished (2003)

Contact: Wm L Crosby

Bioinformatics

University of Saskatchewan, Department of Computer Science

1C101 Engineering Building, 57 Campus Drive, Saskatoon,

Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769

Fax: 306 966 2033

Email: fgas\_escs@usask.ca

This sequence is the direct result of the Base calling software

Phred (default parameters). It is the raw base calls. To aid in the

identification of the high quality insert the software Lucy

(default parameters) has been run on this sequence. Lucy identified

the region [44,808].

Plate: 15B024 row: L column: 10.

Location/Qualifiers

1..1155

FEATURES

source



```

/organism="Triticum aestivum"
/db_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/notes="Vector: pCMV.SPORT6; Crown and developmental stages
of spike formation in wheat cultivar Norstar. 4 mRNA
populations were combined before constructing the library.
The first mRNA population is from 1cm crown sections after
30 days of cold acclimation. The second is from 1cm crown
sections after 11 days of deacclimation (before
deacclimation plants were fully vernalized for 49 days).
The third is from different developmental stages of spike
formation (5 to 50mm) that still have not emerged from the
leaf (dissection required). The last is from different
developmental stages of spike and seed formation after
having emerged from the leaf (visible). First strand
synthesis in this library was done in the presence of
methylated dCTP thereby protecting from internal cleavage
with NotI."

ORIGIN
Alignment Scores:
Pred. No.: 5,09e-91 Length: 1155
Score: 984.50 Matches: 195
Percent Similarity: 82.83% Conservativity: 51
Best Local Similarity: 65.66% Mismatches: 51
Query Match: 35.43% Indels: 3
DB: 7 Gaps: 0

US-10-051-909-36 (1-553) x CK211005 (1-1155)

QY 244 LeuGlyMetPheProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db ::::
54 GTAGGCGATTTGTTTCCTCGAGAAATTCCTCGACATTTTACCTGCTCAATC 113
QY 264 LeuIleProGlyLeuPhePheIleProGlySerProArgTrpLeuAlaLysMetAenLeu 283
Db ::::
114 CTGATACCTGGGTGTTCTTCATCCCGAATCACCAGGTGGCTGGCAAAATGGGAAG 173
QY 284 ThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThr 303
Db ::::
174 ATGGAGGATTTGATCTTCGTCAGGTTCTGGCGGATTTGAACTGATATCTCAGCA 233
QY 304 GluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrIleSerPhe 323
Db ::::
234 GAAGTAAATGAAATTAAGAGATCAGTTGGCTCATCTAGGAGGAGGACACATACCAT 293
QY 324 GlnGluLeuAenGlnLysLysThrArgThrProLeuLeuLeuGlyIleGlyLeuVal 343
Db ::::
294 GCAGAGATCAAAACACAGAGATATAGTGTTCCTTATGATAGGAATTTGCTCCTTATA 353
QY 344 LeuGlnAenLeuSerGlyIleAenGlyValLeuPheThrAlaSerSerIlePheLysAla 363
Db ::::
354 CTTGAGCAACTAGTGGTGTCAATGGCATCTTCTATGCTGCAGTATCTTCAAGCT 413
QY 364 AlaGlyValThrAenSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAla 383
Db ::::
414 GCTGGTCTTAAATAATAGTAATCTAGCAACATGTGTTTGGGGCTGTTCAGTGGTGTCT 473
QY 384 ThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSer 403
Db ::::
474 ACTGGAAATCAACCTGGTGTGACTGACAAAGCTGGTCGACGGCTACTTCTCATTTCTCT 533
QY 404 ThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLysAspAsn 423
Db ::::
534 CTTGTAGGAGATGACAGTCAGTCTCTCTGTTGTTCTGTGTCATTTGTTGAGGAAAC 593
QY 424 IleSerGlnAspSerAenSerTyrrIleLeuThrMetIleSerLeuValGlyIleVal 443
Db ::::
594 ATAGAGGAAGCTTCTCATTTACACTCTGTGATGAGTATGCTTTTCACTGGCTGACTTGTG 653
QY 444 SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlu 463
Db ::::

```

654 GCATTGTGCTTGCATTTCCCTTGGCATGGGAGCCATCCCATGGATCATAATGCTCTGAG 713

464 IleLeuProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAenTrpLeu 483

714 ATCTTCTCTGTTAAATCAAGAGTCTGCGGGAAGCAGCCACCCCTCGCAAACTGGATG 773

484 ThrSerPheAlaIleThrMetThrThrAenLeuMetLeuThrTrpSerValGlyThr 503

774 ACCTCTCTGGTCTATCAGATGACCGGAGCTTGTATCTAACTGGAGCAACGGAGGAACC 833

504 PheLeuSerTyrrMetValValSerAlaPheThrIleValPheValValLeuTrpValPro 523

834 TTGCTATA-TTCGCGCGGTGTCTCATGGGACCCCTCTCTCTGCTGTGCTGTGCGTCCG 892

524 GluThrLysGly\*\*\*AenSerArgGlyAspThrIlePheValSerLeuSer 540

893 GAGACCAAGGAAG-AACATTCAGGAAA-TCGGTTTTCGTTCCGCTGAAC 941

RESULT 8

AV939950 674 bp mRNA linear EST 18-JAN-2002

LOCUS AV939950 K. Sato unpublished cDNA library, strain H602 adult,

DEFINITION heading stage top three leaves Hordeum vulgare subsp. spontaneum

ACCESSION AV939950

VERSION AV939950.1 GI:18235747

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. spontaneum

ORGANISM Hordeum vulgare subsp. spontaneum

REFERENCE 1 (bases 1 to 674)

AUTHORS Sato,K., Saisho,D. and Takeda,K.

TITLE Barley EST sequencing project in NIG and Okayama Univ

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshinigenes.nig.ac.jp.

FEATURES

source

1..674

/organism="Hordeum vulgare subsp. spontaneum"

/mol\_type="mRNA"

/strain="H602"

/sub\_species="spontaneum"

/db\_xref="taxon:77009"

/clone="bah24m12"

/cissue\_type="top three leaves"

/dev\_stage="adult, heading stage"

/clone\_lib="K. Sato unpublished cDNA library, strain H602

adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.: 2,44e-91 Length: 674

Score: 984.00 Matches: 193

Percent Similarity: 91.93% Conservativity: 12

Best Local Similarity: 86.55% Mismatches: 17

Query Match: 35.41% Indels: 1

DB: 1 Gaps: 0

US-10-051-909-36 (1-553) x AV939950 (1-674)

QY 53 GluSerGlySerAspHisAspGlyValLeuArgPro-LeuLeuAenThrGlySerTr 72

Db ::::

3 GAGAGCGGAGCGCACCATGACACCGGNAAGCAGCCGCTGCTGGTCAACACGGGAGCTG 62

QY 72 pTyArgMetSerSerArgGlnSerPheAlaProGlyThrSerSerMetAlaVal 92

Db ::::

[illegible]

```

Qy 312 ValAlaSerSerSerLysArgThrThrIleSerPhe 323
Db 675 GTAGCATCATCAAGGAAGAGACCAACCAATATGTT 710

RESULT 10
AV913420 681 bp mRNA linear EST 18-JAN-2002
LOCUS AV913420 K. Sato unpublished cDNA library, cv. Haruna Nijo
DEFINITION germination shoots Hordeum vulgare subsp. vulgare cDNA clone
bag922a03 5', mRNA sequence.
ACCESSION AV913420
VERSION AV913420.1 GI:18209197
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 681)
AUTHORS Sato, K., Saitoh, D. and Takeda, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..681
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Nijo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="bags22a03"
/tissue_type="shoots"
/dev_stage="germination"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Nijo germination shoots"

ORIGIN
Alignment Scores:
Pred. No.: 2.14e-90 Length: 681
Score: 975.00 Matches: 191
Percent Similarity: 91.89% Conservative: 13
Best Local Similarity: 86.04% Mismatches: 16
Query Match: 35.08% Indels: 2
Gaps: 1

US-10-051-909-36 (1-553) x AV913420 (1-681)

Qy 53 GluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeu---AsnThrGlySer 71
Db 18 GAGGGAGAGAGACCATGAC---ACGCGGAAGCAGCGCTGCTGGTCAACACGGGGAGC 74
Qy 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
Db 75 TGGTACCGGATGGGGTCGCGCAGTCAGCTCCAGCGGGGCGACCTCTCTCATGGCCATC 134
Qy 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProile 111
Db 135 ATGGCGGAGTCCCACGCTCGCCCTTCTCTGCACCATGATCGTCGCGCTCGGCCCATC 194
Qy 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db 195 CAGTTCGGCTTACCGGGGGCTTCTCTCCGCCACCCAGGACGCATCATCCGGGCCTC 254
Qy 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAnValGlyGlyMetVal 151
Db 255 AACCTCTCCATCTCCGAGTTCTCCGTGTTCGGCTCGCTGCCAACGTCGGCGCCATGGTC 314

```

---

```

Qy 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db 315 GGCGCCATCGCCAGCGCCAGATGGCGGAGCAATGGCGGCAAGGGTCACTGATGATT 374
Qy 172 AlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db 375 GCTGCTATTCTTAACATCATCGCTGGCTGGCCATCTCTTCGCAAAAGACACCTCTTT 434
Qy 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 435 CTGTATATGGGACGATTGCTCGAAGGATTGGTGGTGGTGTGCATATCTCATACGGTGCCA 494
Qy 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 495 GTATATACATGACAGAGATTTCTCTCAGAACATGAGAGGGCGCTTAGGCTCTGTGAACAG 554
Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251
Db 555 TTGTCGGTACGATTGGTATCGTGTGGCTACATTCCTCGGCATGTTGTCTCTGGAGG 614
Qy 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db 615 ATGCTTGCAGTAGAGAAATCTTGCATGCACAATATTGATACCTGGTCTGTCTTCTTCA 674
Qy 272 ProGlu 273
Db 675 CCCGAA 680

RESULT 11
LOCUS CR291514 718 bp mRNA linear EST 27-FEB-2004
DEFINITION CR291514 Oryza sativa library (Han B) Oryza sativa cDNA clone
y745e1lp5, mRNA sequence.
ACCESSION CR291514
VERSION CR291514.1 GI:44678080
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 718)
AUTHORS Han, B., Feng, Q., Huang, Y. C., Ying, K., Li, Y., Guan, J. P., Zhu, J. J.,
Zhao, Q. J., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L.,
Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T.,
Zhang, Y. J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X.,
Zhang, L., Lan, L. F., Chen, W., Wu, S. A. and Xue, Y. B.
TITLE Rice cDNA EST clone
JOURNAL Unpublished (2003)
COMMENT Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.
FEATURES
source
1..718
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="y745e1lp5"
/clone_lib="Oryza sativa library (Han B)"

ORIGIN
Alignment Scores:
Pred. No.: 4.62e-89 Length: 718
Score: 962.50 Matches: 198
Percent Similarity: 93.42% Conservative: 15
Best Local Similarity: 86.84% Mismatches: 15
Query Match: 34.63% Indels: 3

```

DB: US-10-051-909-36 (1-553) x CR291514 (1-718) 7 Gaps: 0

160 AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179  
 8 GCGAGCTACATTGGCGCGCAAGGGTCA-TTGTATGATTCTGCAATTCCAAACATCATGTG 66  
 180 TtpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu 199  
 67 TGGCTTGCCATCTCTTTGCCAAAGACATCATCGTTCTTATATGGGACGATGCTCGAG 126  
 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219  
 127 GGGTTTGGTGTGGTGTCTCTTATACGGTGGCGAGTTTACATAGCAGAAATATACCT 186  
 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe 239  
 187 CAAACATCAGAGAGGTCTCTTGGCTCAGGAATCAGTATCTGTACCGTTGGTATATTG 246  
 240 LeuAlaTyrIleLeuGlyMetPheIleProThrArgLeuLeuAlaValIleGlyAlaLeu 259  
 247 TTGGCATATTGCTCGCGCATGTTGTCTTGGAGGATCTTCTGCTGATATAGGAATCTTG 306  
 260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeuAla 279  
 307 CTTTGACACTGTGTGTATGACTGCTCTTATTCATTCAGAAATCCCGAAGATGTTGGCA 366  
 280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr 299  
 367 AAGATGAACATGATGATGATTTTGAGACTCTTTTACAAAGTCTTGAGGGGATTTGAGACT 426  
 300 AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThr 319  
 427 GACATACGCGCGAAGGAATGATATTAAGAGAGCAGTACCTTCAGCAACAAAGAGACA 486  
 320 ThrIleSerPheGlnGluLeuAsnGlnLysGlyTyrArgThrProLeuLeuGlyIle 339  
 487 ACGATCCGTTTCAAGAAATTAACAGCAAGATGAGTATCCGCAACCCCTTAATCTAGGAATT 546  
 340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValIlePheTyrAlaSerSer 359  
 547 GGCCTACTTGTACTGCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 606  
 360 IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379  
 607 ATCTTCCAGCAGCAGGCTCTCAAAACAGTACTTGGGTACATGTGCACCTGGGGGT-ATC 665  
 380 GlnValLeuAlaThrGlyValThr 387  
 666 CAGGTTCT-GCTACAGGAGTTACA 688

RESULT 12  
 AJ794429 762 bp mRNA linear EST 11-AUG-2004  
 LOCUS Antirrhinum majus whole plant Antirrhinum majus cDNA clone  
 DEFINITION 018\_3\_05\_d17, mRNA sequence.  
 ACCESSION AJ794429  
 VERSION AJ794429.1 GI:51109757  
 KEYWORDS EST; Antirrhinum majus (snapdragon)  
 SOURCE Antirrhinum majus  
 ORGANISM Antirrhinum majus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
 asterids; Lamiales; Plantaginaceae; Antirrhineae;  
 Antirrhinum (to 762)  
 Author(s) Zacher, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.  
 Antirrhinum EST collection  
 Antirrhinum (2003)  
 Unpublished (2003)  
 Contact: Schwarz-Sommer Z  
 Molekulare Pflanzen-genetik  
 MPI fuer Zuechtungs-forschung

FEATURES  
 source 1..762  
 Location/Qualifiers  
 /organism="Antirrhinum majus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4151"  
 /clone="018\_3\_05\_d17"  
 /tissue\_type="whole plant"  
 /clone\_lib="Antirrhinum majus whole plant"

ORIGIN  
 Alignment Scores: 5.08e-89 Length: 762  
 Pred. No.: 962.50 Matches: 189  
 Score: 962.50  
 Percent Similarity: 86.00%  
 Best Local Similarity: 75.60%  
 Query Match: 34.63%  
 Indels: 1  
 Gaps: 1  
 DB: 1

US-10-051-909-36 (1-553) x AJ794429 (1-762)

88 SerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAla 107  
 QY |||||  
 DB 13 TCTTCTCAAGCTTTTAGGACAGTTCGTGTTCTGTTGGCTGTGTTGTTGTTGTTGTTGCT 72  
 108 LeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMet 127  
 QY |||||  
 DB 73 TTGGGTCTATCCAGTTTGGTTCACAGTGGTATTCCTCACCGACTCAAACTGCTATT 132  
 128 ValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnVal 147  
 QY |||||  
 DB 133 GTCAATGATCTTAACTCACAGTTCACAGTTCCTGCTATTTGGTCTTTGTTCGAACTTT 192  
 148 GlyGlyMetValGlyAlaIleAlaSerGlyGlyMetAlaGluTyrIleGlyArgLysGly 167  
 QY |||||  
 DB 193 GGAGCGCATGGTTGGAGCAGTAGTACGTTCAGATTCCTCAGTACATAGGAGAGAAAGGGG 252  
 168 SerLeuMetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLys 187  
 QY |||||  
 DB 253 TCTTTAATGATGCGTGCCTACCTACCTAATATCATTTGGTGGCTTGGCTTTCGTTGCCAGA 312  
 188 AspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyPheGlyValGlyIleIleSer 207  
 QY |||||  
 DB 313 GACATCTCAITTTCTGTATCATGGGAGATTTGTTGGAAGGATTTGGCGTCGGCATTAATCTCT 372  
 208 TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly 227  
 QY |||||  
 DB 373 TATACTGTGCTTGTATATAGTACAGTAGACCTGAAATCTAAAGGGGGCTCTCGGG 432  
 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247  
 QY |||||  
 DB 433 TCAGTAACACAGCTCTCTGTGACGATCGGATCATGTAGCATATTTACTTTGGACATATT 492  
 248 IleProTyrArgLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267  
 QY |||||  
 DB 493 TGCATTTGGAGATTTGCTGCTTCTAGGAATATTGCTTGTATATGATATGATACCTGGC 552  
 268 LeuPhePheIleProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCys 287  
 QY |||||  
 DB 553 CTCTTTTTCATCCAGAACTCTCTAGATGTTGGCCAAAATGGGAGATGACAGAGAATTT 612  
 288 GlnThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAsp 307  
 QY |||||  
 DB 613 GAAGCTCTCTTCAAGTTCTTCGGGGGTTTTCGCGTGTATTTGCGATGATGATGATGATGAT 672  
 308 IleLysArgAlaValAlaLysSerSerLysArgThr---ThrIleSerPheGlnGluLeu 326  
 QY |||||  
 DB 673 ATAAAGAAATCTGTTGCTTCAACAGCAGCAGCAGCGGACGAATCCGCTTTTCCCGCATCTT 732  
 327 AsnGlnLysLysTyrArgThrProLeuLeu 336  
 QY |||||  
 DB 733 AAGTTGAAAGACATACCTGGTTACCGTTTGATG 762

```

RESULT 13
CD871121
LOCUS
DEFINITION AZO2.117H06F010207 AZO2 Triticum aestivum cDNA clone AZO2117H06, mRNA sequence.
ACCESSION CD871121
VERSION CD871121.1
KEYWORDS GI:32554937
SOURCE EST.
ORGANISM Triticum aestivum (bread wheat)
Triticum aestivum
Triticum aestivum
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 701)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
source
1. 701
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZO2117H06"
/tissue_type="root"
/clone_lib="AZO2"

ORIGIN
Alignment Scores:
Pred. No.: 6.55e-86 Length: 701
Score: 932.00 Matches: 182
Percent Similarity: 90.48% Conservative: 27
Best Local Similarity: 78.79% Mismatches: 22
Query Match: 33.54% Indels: 0
DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CD871121 (1-701)

Qy 285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGlu 304
Db 7 GAGACTTTGAACCTCTCTACAGTTCTGAGGGGATTCGAGACTGACATCAGCTCAGAA 66

Qy 305 ValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGln 324
Db 67 GTGAATGATATAAGAGAGAGAGAGTAATATACGAAACAAAGGGCGGATCCGTTTCCAA 126

Qy 325 GluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuValLeu 344
Db 127 GAGTTAAACCAAGAAATTCGCGATCCGCTGATCTAGGAATTTGGCCCTGCTTGTCTA 186

Qy 345 GluAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAla 364
Db 187 CAACAGCTAAGCGGAATCAACGCTGATCTGTTATGCAAGTAGCATCTTCAAGCTGCA 246

Qy 365 GlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThr 384
Db 247 GGTATTACAACAGTGAATTTGGCCACATCTGGAGTTCGAGGTATTACAGTTCTTGCAC 306

Qy 385 GlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuIleSerThr 404
Db 307 CTAGTTACAACCTGGTTACTAGACAGGGCTGGCGGCTATCTCTACTCATTATATCTCT 366

Qy 405 SerClyMetThrLeuCysLeuAlaValSerValValPhePheLeuLysAspAsnIle 424
Db 367 GCTGGGATGACTAAGACCTTCTTGGCGTTGCGGTCATATTTTTCAGAGGACATGTT 426

```

```

Qy 425 SerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSer 444
Db 427 TCACAGAGCTCTCACATGATTATACATATTGAGCATGCTCTCCTTGGTTCGTTATTTGGCT 486

Qy 445 PheValIleThrPheSerPheGlyMetGlyAlaIleProTyrIleuMetMetSerGlyIle 464
Db 487 TATGTTATCGCTTCTCCTTCGGTATGGGGGCCATTCATTCGATGATATGCTGAGATT 546

Qy 465 LeuProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAsnThrLeuThr 484
Db 547 CTCGGGTGAGCATCAAGAGTCTCGGGGAAGCTTCGCGACGCTGCCCACTGGCTGACT 606

Qy 485 SerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyIleThrPhe 504
Db 607 TCTTTGGGATACGATGACAGCAAACTTGCCTCAGCTGGAGTCTCGAGGTACATTT 666

Qy 505 LeuSerTyrMetValValSerAlaPheThrIle 515
Db 667 GTGTCATTACATGCTCTGTGAGCGCTTCACGCTC 699

RESULT 14
LOCUS
DEFINITION
CNSO0A8HU
Accession
Version
Keywords
Source
Organism
Reference
Authors
Title
Journal
Reference
Authors
Title
Journal
Comment
Features
Source
1. 1675
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTIS24ZD01"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1. 1675
/gene="At2g48020"

```



Pred. No.: 4.05e-85 Length: 769  
 Score: 925.00 Matches: 177  
 Percent Similarity: 85.66% Conservative: 38  
 Best Local Similarity: 70.52% Mismatches: 36  
 Query Match: 33.29% Indels: 0  
 DB: 4 Gaps: 0

US-10-051-909-36 (1-553) x BI933646 (1-769)

Qy 181 LeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGly 200  
 Db 1 CTTTCCATCTCGTTTGCACAAAGATCTGCTCTTATACATGGAGAGATTATTGGAGGT 60  
 Qy 201 PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln 220  
 Db 61 TTTGGAGTCGGCATAATCTTACACGGTTCTGTATATATGCCGAGATAGCACCTCAG 120  
 Qy 221 AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu 240  
 Db 121 AACCTGAGAGGGCCCTGGGGTCAGTTAACAGCTCTCTGTACATTGGGATCATGTTG 180  
 Qy 241 AlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuPro 260  
 Db 181 GCITATTGTAGGACCTTTTGTAAATTTGGAGAGTGTCTGCTTTCTTGGACATTGCC 240  
 Qy 261 CysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeuAlaLys 280  
 Db 241 TGCTTGCATTGATACCTGCTGCTATTTTTCATCCAGAACTCTCTCGGTGGTTGGCCAAG 300  
 Qy 281 MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAsp 300  
 Db 301 ATGGGTCTGACAGATGATTTTGAACCTCTTTGCAAGTCTCCGAGGGTTCCGATGCTGAC 360  
 Qy 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr 320  
 Db 361 ATTTCCGTCCAGTAATTAATTAAGAGGGCTGTAGCATCCACAGCCGAAAGTCACAA 420  
 Qy 321 IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGly 340  
 Db 421 ATACGTTTTCAGATCTCAAAACAAGAAAGATATTGGCTGCTCTCATGATAGGCATTGGA 480  
 Qy 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle 360  
 Db 481 CTGCTTGTCTACAACTCAACAGCTCAGCGGAACCAATGGTGTGATCTCTTATCCAGTAACAT 540  
 Qy 361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380  
 Db 541 TTCTATCGCCGGGATTTCTTCAAGTATGCTGCACTTTAGGTTTGTGCTATCCAG 600  
 Qy 381 ValLeuAlaThrGlyValThrThrTyrLeuLeuAspArgAlaGlyArgGlyLeuLeu 400  
 Db 601 GTGGTTGCCACTGCTGTTTCTACATGGCTGGTGAATAAACTGGCCGTAGGCTTTTACTG 660  
 Qy 401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu 420  
 Db 661 ATTGCTCTCGCAGCTGGAATGGCTGGTGGTGTGCTCTTATTTTCCATTTCATTGTTG 720  
 Qy 421 LysAspAsnIleSerGlnAspSerAsnSerTyr 431  
 Db 721 AAGGATTTCTAGATGAGGATTTCTACCTCTAT 753

Search completed: April 13, 2005, 18:06:06  
 Job time : 3830.87 secs

*This Page Blank (uspto)*